THEORY OF COMPUTING, Volume 16 (20), 2020, pp. 1–48
www.theoryofcomputing.org

SPECIAL ISSUE: CCC 2019

Fourier and Circulant Matrices are Not Rigid

Zeev Dvir* Allen Liu

Received July 7, 2019; Revised December 18, 2020; Published December 31, 2020

Abstract. The concept of *matrix rigidity* was first introduced by Valiant in 1977. Roughly speaking, a matrix is rigid if its rank cannot be reduced significantly by changing a small number of entries. There has been considerable interest in the explicit construction of rigid matrices as Valiant showed in his MFCS'77 paper that explicit families of rigid matrices can be used to prove lower bounds for arithmetic circuits.

In a surprising recent result, Alman and Williams (FOCS'19) showed that the $2^n \times 2^n$ Walsh–Hadamard matrix, which was conjectured to be rigid, is actually not very rigid. This line of work was extended by Dvir and Edelman (*Theory of Computing*, 2019) to a family of matrices related to the Walsh–Hadamard matrix, but over finite fields. In the present paper we take another step in this direction and show that for any abelian group G and function $f : G \to \mathbb{C}$, the *G-circulant matrix,* given by $M_{xy} = f(x-y)$ for $x, y \in G$, is not rigid over $\mathbb{C}$. Our results also hold if we replace $\mathbb{C}$ with a finite field $\mathbb{F}_q$ and require that $\gcd(q, |G|) = 1$. En route to our main result, we show that circulant and Toeplitz matrices (over finite fields or $\mathbb{C}$) and Discrete Fourier Transform (DFT) matrices (over $\mathbb{C}$) are not sufficiently rigid to carry out Valiant's approach to proving circuit lower bounds. This complements a recent result of Goldreich and Tal (*Comp. Complexity*, 2018) who showed that Toeplitz matrices are nontrivially rigid (but not enough for Valiant's method). Our work differs from previous

A preliminary version of this paper appeared in the Proceedings of the 34th IEEE Conference on Computational Complexity, 2019.

*Research supported by NSF CAREER award DMS-1451191 and NSF grant CCF-1523816.

ACM Classification: F.2.2, F.1.3

AMS Classification: 68Q17, 68Q15

Key words and phrases: matrix rigidity, circulant matrix

 DOI: 10.4086/toc.2020.v016a020

non-rigidity results in that those papers considered matrices whose underlying group of symmetries was of the form $\mathbb{Z}_p^n$ with p fixed and n tending to infinity, while in the families of matrices we study, the underlying group of symmetries can be any abelian group and, in particular, the cyclic group $\mathbb{Z}_N$, which has very different structure. Our results also suggest natural new candidates for rigidity in the form of matrices whose symmetry groups are highly non-abelian.

Contents

1 Introduction

1.1 Background

A major goal in complexity theory is to prove lower bounds on the size and depth of arithmetic circuits that compute certain functions. One specific problem that remains open despite decades of effort is to find functions for which we can show super-linear size lower bounds for circuits of logarithmic depth. In [19], Valiant introduced the notion of matrix rigidity as a possible method of proving such lower bounds for arithmetic circuits. More precisely, over a field $\mathbb{F}$, an $m \times n$ matrix M is said to be (r,s)-rigid if any $m \times n$ matrix of rank at most r differs from M in at least s entries. Valiant showed that for any linear function $f : \mathbb{F}^n \to \mathbb{F}^n$ that can be computed by an arithmetic circuit of size $O(n)$ and depth $O(\log n)$, the corresponding matrix can be reduced to rank $O(\frac{n}{\log\log n})$ by changing $O(n^{1+\varepsilon})$ entries for any $\varepsilon > 0$. Thus, to prove a circuit lower bound for a function f, it suffices to lower bound the rigidity of the corresponding matrix at rank $O(\frac{n}{\log\log n})$. We call a matrix Valiant-rigid if it is $\left(O(\frac{n}{\log\log n}), \Omega(n^{1+\varepsilon})\right)$-rigid for some $\varepsilon > 0$, i. e., sufficiently rigid for Valiant's method to yield circuit lower bounds. Over any infinite field, Valiant shows that almost all $n \times n$ matrices are $(r,(n-r)^2)$-rigid for any r, while over a finite field one can get a similar result with a logarithmic loss in the sparsity parameter. Despite much effort, explicit constructions of rigid matrices have remained elusive.

Over infinite (or very large) fields, there are ways to construct highly rigid matrices using either algebraically independent entries or entries that have exponentially large description (see [16, 12, 15]).[1] However, these constructions are not considered to be fully explicit as they do not tell us anything about the computational complexity of the corresponding function. Ideally, we would be able to construct rigid $(0,1)$-matrices, but even a construction where the entries are in a reasonably simple field (such as the m^{th} cyclotomic field for a small value of m) would be a major breakthrough. The best known constructions of such matrices are $(r, \Omega(\frac{n^2}{r}\log\frac{n}{r}))$-rigid (see [18, 9]). There has also been work towards constructing semi-explicit rigid matrices. Semi-explicit constructions which require $O(n)$ bits of randomness (instead of the usual $O(n^2)$) would still yield circuit lower bounds through Valiant's approach.[2] The best result in this realm (see [11]) shows that random Toeplitz matrices are $(r, \frac{n^3}{r^2\log n})$-rigid with high probability for $r \geq \Omega(\sqrt{n})$.

Note that both of these bounds become trivial when r is $n/\log\log n$. Other variants of semi-explicit constructions have also been studied. [1] gives a construction of $(2^{(\log n)^{1/4-o(1)}}, \Omega(n^2))$-rigid matrices using an NP-oracle. This construction is not in the regime for Valiant-rigidity.

Many well-known families of matrices, such as Hadamard matrices (square matrices with ± 1 entries whose rows are orthogonal) and DFT (Discrete Fourier Transform) matrices, have been conjectured to be Valiant-rigid [17]. However, a recent line of work (see [2, 7]) shows that certain well-structured matrices are not rigid. Alman and Williams show in [2] that the Walsh–Hadamard matrix, i. e., the $2^n \times 2^n$ Hadamard matrix given by $H_{xy} = (-1)^{\langle x,y\rangle}$ as x and y range over $\{0,1\}^n$, is not Valiant-rigid over $\mathbb{Q}$. Along similar lines, Dvir and Edelman show in [7] that G-circulant matrices for the additive group of $\mathbb{F}_p^n$,

[1] It remains open to construct a matrix that is Valiant-rigid, even if we only require that the entries live in a number field of dimension polynomial in the size of the matrix.

[2] Note however, that it is easy to construct rigid matrices with $O(n^{1+\varepsilon})$ bits of randomness for any $\varepsilon > 0$ (for example by taking a random matrix with at most n^{ε} non-zeros per row) but this is not sufficient for Valiant's approach.

given by $M_{xy} = f(x-y)$ where $f : \mathbb{F}_p^n \to \mathbb{F}_p$ and x, y range over $\mathbb{F}_p^n$, are not Valiant-rigid over $\mathbb{F}_p$ (where we view p as fixed and n goes to infinity). The Walsh–Hadamard matrix and the G-circulant matrices for the additive group of $\mathbb{F}_p^n$ have the property that for any $\varepsilon > 0$, there exists an $\varepsilon' > 0$ such that it is possible to change at most $N^{1+\varepsilon}$ entries and reduce the rank to $N^{1-\varepsilon'}$ (where N denotes the size of the matrix). The proofs of both results rely on constructing a matrix determined by a polynomial $P(x, y)$ that agrees with the given matrix on almost all entries and then arguing that the constructed matrix has low rank.

1.2 Our contribution

Definition 1.1 (G-circulant matrices)**.** Let G be a finite abelian group, $\mathbb{F}$ a field, and $f : G \to \mathbb{F}$ a function. The G-circulant matrix $M(f)$ is defined as the $|G| \times |G|$ matrix whose rows and columns are labeled by the elements of G and whose (x, y) entry is $f(x-y)$ (for $x, y \in G$).

In this paper we prove that for an abelian group G, over any finite field with characteristic relatively prime to $|G|$ and over the complex numbers, G-circulant matrices are not Valiant-rigid (Theorem 1.5). En route to our main result, we prove that the following commonly studied families of matrices are not Valiant-rigid:

- DFT matrices (over $\mathbb{C}$).
- Circulant matrices (over finite fields and $\mathbb{C}$): matrices whose rows are obtained by cyclically shifting the top row.
- Toeplitz matrices (over finite fields and $\mathbb{C}$): matrices with constant diagonals.[3]

Remark 1.2. For circulant and Toeplitz matrices over finite fields, we do not require any additional conditions, i.e. we do not require that the size of the matrix and the characteristic of the field are relatively prime. See the beginning of Section 8 for a more in-depth discussion about why we require such a condition for general abelian groups.

The families of matrices we consider in our paper have very different underlying group structure than those considered in previous work. Both [2] and [7] analyze matrices constructed from an underlying group of the form $\mathbb{Z}_p^n$ with p a fixed prime number and n tending to infinity. In this paper we study matrices whose underlying symmetry group can be any abelian group. In fact, the core of our proof is handling the case when the underlying group is cyclic.

Circulant matrices are the special case of G-circulants for cyclic groups G. Similarly, the DFT matrices are the special case of the DFT_G matrices (where DFT_G is the matrix given by the character table of an abelian group G) when G is cyclic. The Walsh–Hadamard matrices are another special case of the DFT_G matrices, where G is the group $\mathbb{Z}_2^n$. We use the fact, that every finite abelian group can be decomposed into the direct product of cyclic groups, to extend our results to all abelian groups, although this extension is by no means immediate. While most natural constructions of matrices are highly symmetric, our results show that matrices that are symmetric under abelian groups are not rigid

[3]It is not hard to see that rigidity of circulant and Toeplitz matrices is essentially the same question so for the sake of consistency with our (group theoretic) approach we will primarily consider circulant matrices.

and that perhaps we should look toward less structured matrices, or matrices whose symmetry group is non-abelian, as candidates for rigidity.

We now move into a more technical overview of our paper. We begin with a few definitions.

Definition 1.3. Define the *regular-rigidity* $\mathsf{r}_A^{\mathbb{F}}(r)$ of a matrix A over a field $\mathbb{F}$ as the minimum value of s such that it is possible to change at most s entries in each row and column of A to obtain a matrix of rank at most r.

When the field is clear from context, we will omit the superscript.The notion of regular-rigidity is weaker than the usual notion of rigidity (and is also weaker than the commonly used notion of row-rigidity) as if A is an $n \times n$ matrix and A is (r, ns)-rigid then $\mathsf{r}_A(r) \geq s$. Note that this actually makes our results stronger as we will show that the matrices we consider are not regular-rigid.

To simplify the exposition, we define a qualitative notion of non-rigidity we call QNR (quasipolynomial non-rigidity).

Definition 1.4. We say that a family $\mathcal{A}$ of matrices is *quasipolynomially non-rigid* (QNR) over a field $\mathbb{F}$ if there are constants $c_1, c_2 > 0$ such that for any $\varepsilon > 0$, all sufficiently large matrices $M \in \mathcal{A}$ satisfy

$$\mathsf{r}_M^{\mathbb{F}}\left(\frac{N}{\exp\left(\varepsilon^{c_1}(\log N)^{c_2}\right)}\right) \leq N^{\varepsilon},$$

where M is an $N \times N$ matrix.

We will prove that various families of matrices are QNR. Note that this immediately implies that they are not Valiant-rigid. Our main results are stated below.

Theorem 1.5. *Let G be an abelian group. The family of G-circulant matrices is QNR over $\mathbb{C}$. For a finite field $\mathbb{F}_q$, if $\gcd(|G|, q) = 1$, then the family of G-circulant matrices is QNR over $\mathbb{F}_q$.*

The formal statement and proof of this result can be found in Section 6 (see Theorem 6.2) for the complex numbers and Section 8 (see Theorem 8.5) for finite fields.

Theorem 1.6. *Let G be an abelian group of order N. Then there exists $m = \tilde{O}(N^3)$, depending only on G, such that the rational G-circulant matrices are QNR over the m^{th} cyclotomic field. The same also holds for the matrix DFT_G.*

The formal statement and proof of this result can be found in Section 6 (see Theorem 6.3).

In addition to the aforementioned results for circulant, Toeplitz and DFT matrices, our main theorem has a few more consequences that are worth mentioning. The following two corollaries to our main result were pointed out by Babai and Kivva [4]:

- The Paley–Hadamard matrices are QNR over $\mathbb{C}$.
- The Vandermonde matrices $V_n(x_1, \ldots, x_n)$ whose generators $x_1, \ldots, x_n$ form a geometric progression are QNR over $\mathbb{C}$.

1.3 Overview of the proof

We now take a more detailed look at the techniques used in the proof of Theorem 1.5.

In general, matrices that we deal with will be over $\mathbb{C}$ except in Sections 7 and 8 where we extend our results to matrices over finite fields. First we define two families of matrices that we will use extensively.

Definition 1.7 (Generalized Walsh–Hadamard (GWH) matrices)**.** The generalized Walsh–Hadamard (GWH) matrix $H_{d,n}$ is a $d^n \times d^n$ complex matrix that has rows and columns indexed by $\mathbb{Z}_d^n$ and entries $(H_{d,n})_{I,J} = \omega^{I \cdot J}$ where $\omega = \mathrm{e}^{2\pi i/d}$.

Next we define the Discrete Fourier Transform (DFT) matrices.

Definition 1.8 (DFT matrix)**.** The (x,y) entry of the $N \times N$ matrix DFT_N $(0 \leq x, y \leq N-1)$ is ω^{xy} where $\omega = \mathrm{e}^{2\pi i/N}$.

Note $\mathrm{DFT}_N = H_{N,1}$ and $H_{d,n} = \underbrace{\mathrm{DFT}_d \otimes \cdots \otimes \mathrm{DFT}_d}_{n}$ where $\otimes$ denotes the Kronecker product.

One key idea in our argument is the observation that, if all members of a family $\mathcal{A}$ of matrices are simultaneously diagonalizable by a matrix M, then the rigidity of *any* matrix $A \in \mathcal{A}$ implies the rigidity of the matrix M (Lemma 2.21). This situation happens, e. g., when $\mathcal{A}$ is the family of circulant matrices and M is the DFT matrix. This simple, yet crucial observation allows us to deduce the non-rigidity of a larger family of matrices.[4]

1.3.1 Generalized Walsh–Hadamard matrices

The first step in the proof of Theorem 1.5 is proving that the generalized Walsh–Hadamard matrices are not rigid in the following sense, which is stronger than QNR.

Theorem 1.9 (Generalized Walsh–Hadamard matrices are not rigid)**.** *For fixed d and $0 < \varepsilon < 0.01$, there exists an ε' such that for all sufficiently large n, $r_{H_{d,n}}\left(d^{n(1-\varepsilon')}\right) \leq d^{n\varepsilon}$.*

Note that Theorem 1.9 generalizes the main result of [2] (which deals with $d = 2$). The result of [2] is stronger in the sense that it holds over $\mathbb{Q}$ while our results for GWH matrices require working over a field extension. See Section 1.4 for a discussion on rigidity over different fields.

Also, given any $d^n \times d^n$ matrix of the form $M_{xy} = f(x-y)$ with $f : \mathbb{Z}_d^n \to \mathbb{C}$, we can permute its rows so that it is diagonalized by $H_{d,n}$. Thus, we can apply the diagonalization trick mentioned above and obtain the following result, which extends the results in [7] to matrices over $\mathbb{C}$.

Corollary 1.10. *Let f be a function from $\mathbb{Z}_d^n \to \mathbb{C}$ and let M be a $d^n \times d^n$ matrix with $M_{xy} = f(x-y)$. Then for any fixed d and $0 < \varepsilon < 0.01$, there exists an $\varepsilon' > 0$ such that for all sufficiently large n, we have $r_M\left(d^{n(1-\varepsilon')}\right) \leq d^{n\varepsilon}$.*

[4]The observation that DFT matrices diagonalize circulant matrices has similar algorithmic implications as if there were, say, linear-size circuits for computing the DFT matrix then we would be able to obtain linear-size circuits for computing any convolution.

1.3.2 DFT matrices

Equipped with the machinery for Generalized Walsh–Hadamard (GWH) matrices, the next step is to prove non-rigidity for DFT matrices. The result we prove is the following.

Theorem 1.11 (DFT Matrices are Not Rigid)**.** *The family of DFT matrices* DFT_N *where* $N \in \mathbb{N}$ *is QNR over* $\mathbb{C}$.

Our proof consists of two steps. First we show that for integers N of a very special form, the $N \times N$ DFT matrix is not rigid because it can be decomposed into submatrices with GWH-type structure. We say an integer N is *well-factorable* if it is a product of distinct primes $q_1, \dots, q_l$ such that $q_i - 1$ has no large prime power divisors for all i. We will make this notion more precise later, but informally, the first step is as follows.

Theorem 1.12. *Let* $\mathcal{A}$ *denote the family of DFT matrices* DFT_N *where* N *is well-factorable. Then the family* $\mathcal{A}$ *is QNR over* $\mathbb{C}$.

The main intuition is that if N is a product of distinct primes $q_1, \dots, q_l$, then within the DFT matrix DFT_N, we can find submatrices whose rows and columns can be indexed by $\mathbb{F}_{q_1}^\times \times \cdots \times \mathbb{F}_{q_l}^\times$ where $\mathbb{F}^\times$ denotes the multiplicative group of the field $\mathbb{F}$. This multiplicative structure can be replaced by the additive structure of $\mathbb{Z}_{q_1-1} \times \cdots \times \mathbb{Z}_{q_l-1}$. We can then factor each additive group $\mathbb{Z}_{q_i-1}$ into prime power components. If $q_1 - 1, \dots, q_l - 1$ all have no large prime power divisors, we expect prime powers to be repeated many times when all of the terms are factored. This allows us to find submatrices with $\mathbb{Z}_d^l$ additive structure to which we can apply tools such as Theorem 1.9 and Corollary 1.10 to reduce the rank while changing a small number of entries. We then bound the rank and total number of entries changed over all submatrices to deduce that DFT_N is not rigid.

The second step of our proof that DFT matrices are not rigid involves extending Theorem 1.12 to all values of N. The diagonalization trick gives that $N \times N$ circulant matrices are not rigid when N is well-factorable. We then show that for $N' < \frac{N}{2}$, we can rescale the columns of the $N' \times N'$ DFT matrix and embed it into an $N \times N$ circulant matrix. As long as N' is not too much smaller than N (say $N' > \frac{N}{(\log N)^2}$), we get that the $N' \times N'$ DFT matrix is not rigid. Thus, for each well-factorable N and all N' in the range $\frac{N}{(\log N)^2} < N' < \frac{N}{2}$, the $N' \times N'$ DFT matrix is not rigid. We then use a number theoretic result of Baker and Harman [5] to show that the multiplicative gaps between well-factorable integers are not too large. Thus, the above intervals cover all integers as N runs over all well-factorable numbers, finishing the proof.

As a corollary to Theorem 1.11 (due to the diagonalization trick), we get that circulant matrices are not rigid.

Theorem 1.13 (Circulant Matrices are not Rigid)**.** *Let* $\mathcal{A}$ *denote the family of circulant matrices. Then* $\mathcal{A}$ *is QNR over* $\mathbb{C}$.

Also notice that since any Toeplitz matrix of size at most $\frac{N}{2}$ can be embedded in an $N \times N$ circulant matrix, the above implies an analogous result for all Toeplitz matrices. While [11] shows nontrivial rigidity lower bounds for rank much smaller than N, our results imply that there are actually no nontrivial rigidity lower bounds for rank close to N.

1.3.3 G-circulant matrices

We extend our results for DFT and circulant matrices to DFT_G and G-circulant matrices for finite abelian groups G by using the fundamental theorem of finite abelian groups to write G as a direct product of cyclic groups. Note that any G-circulant matrix is diagonalized by the DFT_G matrix which is the Kronecker product of the DFT matrices for the individual cyclic groups. If there are many small cyclic groups in the product, then we can use the same techniques that we use for GWH-matrices while if there are enough large cyclic groups, then we can rely on our results for DFT matrices.

1.4 Rigidity over different fields

There are several interesting questions that arise when considering rigidity over different fields. Our results for circulant and DFT matrices require working over a field extension. The matrix DFT_N is defined over the N-th cyclotomic field $\mathbb{Q}[\omega]$ where ω is a primitive N^{th} root of unity. But we are not able to show non-rigidity of this matrix over $\mathbb{Q}[\omega]$, only over a larger field that includes additional roots of unity. Therefore the same holds for circulant matrices over $\mathbb{Q}$ whose non-rigidity we derived from the non-rigidity of DFT_N. The degree of the extension is $\tilde{O}(N^2)$ (so combined with ω, the entire extension of $\mathbb{Q}$ has degree $\tilde{O}(N^3)$). See Theorem 6.3 for more details.

We leave it as an open question whether our results for fields of characteristic 0 still hold without field extensions or for extensions of lower degree.

In Section 7 we extend our results for complex-valued matrices to any finite field, $\mathbb{F}_q$. The main idea is to first work over an extension, say $\mathbb{F}_q[\alpha]$, where the matrices are not rigid, and then sum over the matrices obtained by replacing α with its conjugates. A key ingredient in our proof is that over finite fields, primitive n^{th} roots of unity may have minimal polynomial with very low degree, even subpolynomial in n. However, over $\mathbb{Q}$, the primitive n^{th} roots of unity have minimal polynomial with degree $\phi(n)$ (which is $n^{1-o(1)}$) so our argument does not generalize to this case. We leave it as an open question whether circulant matrices are rigid over $\mathbb{Q}$. It is worth noting that the result of Alman and Williams in [2] does hold over $\mathbb{Q}$ while the result of Dvir and Edelman in [7] holds only when the field is a finite field related to the group structure of the matrix.

Finally, over a finite field $\mathbb{F}_q$, our result for G-circulant matrices requires that $\gcd(q,|G|)=1$. Our result for circulant matrices (i.e. cyclic groups) over finite fields does not require this assumption. The reason that we need $\gcd(q,|G|)=1$ for general abelian groups is that our techniques do not deal with groups such as $G=\mathbb{Z}_{p^2}\times\cdots\times\mathbb{Z}_{p^2}$ (where p is the characteristic of $\mathbb{F}_q$) because p^{th} roots of unity do not exist over any extension of $\mathbb{F}_q$ so we cannot diagonalize G-circulant matrices even if we lift to a field extension. This is not an issue for large cyclic groups because for a cyclic group, say $\mathbb{Z}_N$, we can embed a $\mathbb{Z}_N$-circulant matrix in a circulant matrix of any given size at least $2N$. We only require the condition $\gcd(q,|G|)=1$ to rule out the case when G contains a direct product of many copies of the same small cyclic group whose order is not relatively prime to q. See Section 8 for more details. We leave it as an open question whether our results for G-circulant matrices still hold without the condition that $\gcd(q,|G|)=1$. The results of Dvir and Edelman in [7] deal with a special case where $\gcd(q,|G|)>1$, namely when G is a direct product of many copies of $\mathbb{Z}_p$ where p is the characteristic of $\mathbb{F}_q$.

1.5 Organization

In Section 2, we introduce notation and prove several basic results that we will use throughout the paper. In Section 3, we show that GWH matrices and several closely related families of matrices are not rigid. In Section 4, we show that $N \times N$ DFT matrices are not rigid when N satisfies certain number-theoretic conditions. In Section 5, we complete the proof that no (sufficiently large) DFT matrix is rigid. We then deduce that Toeplitz matrices are not rigid. In Section 6, we use the results from the previous section to show that G-circulant matrices for abelian groups G are not rigid. From Section 2 through Section 6, we work with matrices over $\mathbb{C}$ for ease of exposition. In Section 7 and Section 8, we sketch how to modify the proofs in the previous sections to deal with "missing" roots of unity in a finite field. Finally, in Section 9, we discuss a few open questions and possible directions for future work.

2 Preliminaries

Throughout this paper, we let $d \geq 2$ be an integer and $\omega = \mathrm{e}^{2\pi i/d}$ be a primitive d^{th} root of unity. When we consider an element of $\mathbb{Z}_d^n$, we will view it as an ordered n-tuple with entries in the range $[0, d-1]$. When we say a list of d^n elements $x_1, \ldots, x_{d^n}$ is indexed by $\mathbb{Z}_d^n$, we mean that each x_i is labeled with an element of $\mathbb{Z}_d^n$ such that all labels are distinct and the labels of $x_1, \ldots, x_{d^n}$ are in lexicographical order.

2.1 Basic notation

We will frequently work with ordered tuples, say $I = (i_1, \ldots, i_n) \in \mathbb{Z}_d^n$. Below we introduce some notation for dealing with ordered tuples that will be used later on.

Definition 2.1. For an ordered tuple I, we let $I^{(i)}$ denote its i^{th} entry. For instance if $I = (i_1, \ldots, i_n)$ then $I^{(k)} = i_k$.

Definition 2.2. For an ordered n-tuple $I = (i_1, i_2, \ldots, i_n)$, define the polynomial in n variables $x^I = x_1^{i_1} \cdots x_n^{i_n}$.

Definition 2.3. For ω a d^{th} root of unity and an ordered n-tuple $I = (i_1, i_2, \ldots, i_n) \in \mathbb{Z}_d^n$, we define $\omega^{[I]} = (\omega^{i_1}, \ldots, \omega^{i_n})$.

Definition 2.4. For a function $f : \mathbb{Z}_d^n \to \mathbb{C}$, define the n-variable polynomial P_f as

$$P_f = \sum_{I \in \mathbb{Z}_d^n} f(I) x^I \,.$$

Definition 2.5. For an ordered n-tuple $I = (i_1, i_2, \ldots, i_n)$, we define the set $\mathsf{perm}(I)$ to be a set of ordered n-tuples consisting of all distinct permutations of the entries of I. Similarly, for a set of ordered n-tuples S, we define $\mathsf{perm}(S)$ to be the set of all ordered n-tuples that can be obtained by permuting the entries of some element of S.

Definition 2.6. We say a set $S \subseteq \mathbb{Z}_d^n$ is symmetric if $\mathsf{perm}(I) \subseteq S$ for any $I \in S$.

Definition 2.7. For a set of ordered n-tuples S, let $\mathsf{red}(S)$ denote the set of equivalence classes under permutation of entries in S. Let $\mathsf{rep}(S)$ be a set of ordered n-tuples formed by taking one representative from each equivalence class in $\mathsf{red}(S)$ (note $\mathsf{rep}(S)$ is not uniquely determined but this will not matter for our purposes).

Note that if $\mathsf{rep}(S) = \{I_1, \dots, I_k\}$, then the sets $\mathsf{perm}(I_1), \mathsf{perm}(I_2), \dots, \mathsf{perm}(I_k)$ are disjoint and their union contains S. If the set S is symmetric then their union is exactly S.

2.2 Special families of matrices

We now define notation for working with a few special families of matrices.

Definition 2.8. An $N \times N$ matrix M is called a Toeplitz matrix if M_{ij} depends only on $i - j$. An $N \times N$ matrix M is called a Hankel matrix if M_{ij} depends only on $i + j$. Note that the rows of any Toeplitz matrix can be permuted to obtain a Hankel matrix so any non-rigidity results we show for one family also hold for the other.

Definition 2.9 (Adjusted G-circulant matrices)**.** For an abelian group G and a function $f : G \to \mathbb{C}$, let $M_G(f)$ denote the $|G| \times |G|$ matrix (over $\mathbb{C}$) whose rows and columns are indexed by elements $x, y \in G$ and whose entries are given by $M_{xy} = f(x+y)$. When it is clear what G is from context, we will simply write $M(f)$. We let V_G denote the family of matrices $M_G(f)$ as f ranges over all functions from G to $\mathbb{C}$. We call V_G the family of *adjusted G-circulant matrices* for the group G. When G is a cyclic group, we call the matrices in V_G adjusted-circulant.

Compared to the usual G-circulant (and circulant) matrices defined by $M_{xy} = f(x-y)$, the matrix $M_G(f)$ differs only in a permutation of the rows. In the subsequent sections, we will work with $M_G(f)$ for technical reasons, but it is clear that the same non-rigidity results hold for the usual G-circulant matrices. Similarly, we will use adjusted-circulant and Hankel matrices as it is clear that the same non-rigidity results hold for circulant and Toeplitz matrices. Also note that adjusted-circulant matrices are a special case of Hankel matrices.

Recall that a *character* of an abelian group G is a homomorphism from G to $\mathbb{C}^{\times}$, the multiplicative group of complex numbers.

Definition 2.10 (Discrete Fourier Transfrom matrices)**.** For a finite abelian group G, we define DFT_G, the DFT matrix for G, as the $|G| \times |G|$ matrix whose rows correspond to elements of G and whose columns correspond to the characters of the group. To simplify notation, we will write DFT_N for $\mathrm{DFT}_{\mathbb{Z}_N}$, the classical $N \times N$ Fourier Transform matrix (for the cyclic group $\mathbb{Z}_N$).

The following is immediate from the definition.

Fact 2.11. For a finite abelian group G, if $G = H \times K$ where H and K are subgroups, there is an ordering of the rows and columns of G so that $\mathrm{DFT}_G = \mathrm{DFT}_H \otimes \mathrm{DFT}_K$. In particular, if $G = \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_a}$ then

$$\mathrm{DFT}_G = \mathrm{DFT}_{n_1} \otimes \cdots \otimes \mathrm{DFT}_{n_a} .$$

2.3 Matrix rigidity

Here, we review basic notation for matrix rigidity.

Definition 2.12. For a matrix M and a real number r, we define $\mathsf{R}_M(r)$ to be the smallest number s for which there exists a matrix A with at most s nonzero entries and a matrix B of rank at most r such that $M = A + B$. If $\mathsf{R}_M(r) \geq s$, we say M is (r,s)-rigid.

Definition 2.13. For a matrix M and a real number r, we define $\mathsf{r}_M(r)$ to be the smallest number s for which there exists a matrix A with at most s nonzero entries in each row and column and a matrix B of rank at most r such that $M = A + B$. If $\mathsf{r}_M(r) \geq s$, we say M is (r,s)-regular rigid.

It is clear that if a matrix is (r,ns)-rigid, then it must be (r,s)-regular rigid. In the following sections, we will show that various matrices are not $(\frac{N}{\log\log N}, N^{\varepsilon})$-regular rigid for any $\varepsilon > 0$ and this will imply that Valiant's method for showing circuit lower bounds in [19] cannot be applied for these matrices.

2.4 Preliminary results

Next, we mention several basic results that will be useful in the proofs later on.

Definition 2.14. For an $m \times n$ matrix A and $p \times q$ matrix B, the Kronecker product $A \otimes B$ is the $mp \times nq$ matrix given by

$$\begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix}$$

where a_{ij} are the entries of A.

Fact 2.15. For matrices A,B,C,D such that matrix products AC and BD are defined,

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)\,.$$

Claim 2.16. *$H_{d,n} = \underbrace{\mathrm{DFT}_d \otimes \cdots \otimes \mathrm{DFT}_d}_{n}$ where $\otimes$ denotes the Kronecker product.*

Proof. This can easily be verified from the definition. □

Claim 2.17. *$H_{d,n}H_{d,n}^* = d^n I$ where $H_{d,n}^*$ is the conjugate transpose of $H_{d,n}$ and I is the identity matrix.*

Proof. We verify that $\mathrm{DFT}_d\,\mathrm{DFT}_d^* = dI$, and then use Claim 2.17 and Fact 2.15. □

Claim 2.18. *Let $f : \mathbb{Z}_d^n \to \mathbb{C}$ be a function. Let ω be a primitive d^{th} root of unity and set $P_f = \sum_{I \in \mathbb{Z}_d^n} f(I) x^I$ (see Definition 2.2). Let $D = H_{d,n} M_{\mathbb{Z}_d^n}(f) H_{d,n}$. Then D is a diagonal matrix with diagonal entries $d^n P_f(\omega^{[J]})$ as J ranges over $\mathbb{Z}_d^n$.*

Proof. First, we analyze the product $M_{\mathbb{Z}_d^n}(f)H_{d,n}$. This is a $d^n \times d^n$ matrix and its rows and columns can naturally be indexed by ordered tuples $I, J \in \mathbb{Z}_d^n$. The entry with row indexed by I and column indexed by J is

$$\sum_{I' \in \mathbb{Z}_d^n} f(I+I')\omega^{I' \cdot J} = \omega^{-I \cdot J} \sum_{I' \in \mathbb{Z}_d^n} f(I+I')\omega^{(I'+I)\cdot J} = \omega^{-I \cdot J} P_f(\omega^{[J]}).$$

Therefore, the columns of $M_{\mathbb{Z}_d^n}(f)H_{d,n}$ are multiples of the columns of $H^*_{d,n}$. In fact, the column of $M_{\mathbb{Z}_d^n}(f)H_{d,n}$ indexed by J is $P_f(\omega^{[J]})$ times the corresponding column of $H^*_{d,n}$. Since $H_{d,n}H^*_{d,n} = d^n I$, we deduce that D must be a diagonal matrix whose entries on the diagonal are $d^n P_f(\omega^{[J]})$ as J ranges over $\mathbb{Z}_d^n$. □

Claim 2.19. *Let M be a $d \times d$ adjusted-circulant matrix. Then $\mathrm{DFT}_d \cdot M \cdot \mathrm{DFT}_d$ is a diagonal matrix.*

Proof. Plug $n = 1$ into the above. □

Claim 2.18 gives us a characterization of the rank of matrices of the form $M_{\mathbb{Z}_d^n}(f)$.

Claim 2.20. *Let $f : \mathbb{Z}_d^n \to \mathbb{C}$ be a function. Let ω be a d^{th} root of unity and assume $P_f = \sum_{I \in \mathbb{Z}_d^n} f(I)x^I$ has C roots among the set $\{(\omega^{i_1}, \ldots, \omega^{i_n}) \mid (i_1, \ldots, i_n) \in \mathbb{Z}_d^n\}$. Then $\mathsf{rank}(M_{\mathbb{Z}_d^n}(f)) = d^n - C$.*

Proof. Consider the product $D = H_{d,n} M_{\mathbb{Z}_d^n}(f) H_{d,n}$. Note that $H_{d,n}$ is clearly invertible by Claim 2.17. Therefore, it suffices to compute the rank of D. By Claim 2.18, D must be a diagonal matrix whose entries on the diagonal are $d^n P_f(\omega^{[J]})$ as J ranges over $\mathbb{Z}_d^n$. The rank of D is the number of nonzero diagonal entries which is simply $d^n - C$. □

As mentioned in the introduction, we can relate the rigidity of a matrix to the rigidity of matrices that it diagonalizes.

Lemma 2.21. *If $B = A^*DA$ where D is a diagonal matrix and $\mathsf{r}_A(r) \le s$ then $\mathsf{r}_B(2r) \le s^2$. The same inequality holds also for $B' = ADA$.*

Proof. Let E be the matrix with at most s nonzero entries in each row and column such that $A - E$ has rank at most r. We have

$$B - E^*DE = A^*D(A-E) + (A^* - E^*)DE.$$

Since $\mathsf{rank}(A-E) \le r$, we get that $\mathsf{rank}(B - E^*DE) \le 2r$. Also, E^*DE has at most s^2 nonzero entries in each row and column so $\mathsf{r}_B(2r) \le s^2$. The second part can be proved in the exact same way with A^* replaced by A. □

In light of Lemma 2.21, Claim 2.19, and Claim 2.18, proving non-rigidity for $d \times d$ circulant matrices reduces to proving non-rigidty for DFT_d and proving non-rigidity for G-circulant matrices for $G = \mathbb{Z}_d^n$ reduces to proving non-rigidity for $H_{d,n}$. Below, we show that these statements are actually equivalent.

Claim 2.22. *It is possible to rescale the rows and columns of $H_{d,n}$ to get a matrix of the form $M_{\mathbb{Z}_d^n}(f)$ for some symmetric function $f : \mathbb{Z}_d^n \to \mathbb{C}$. In particular, it is possible to rescale the rows and columns of DFT_d to get an adjusted-circulant matrix.*

Proof. Let ζ be such that $\zeta^2 = \omega$. Multiply each row of $H_{d,n}$ by $\zeta^{(I\cdot I)}$ and each column by $\zeta^{(J\cdot J)}$ to get a matrix H'. We have

$$H'_{IJ} = \zeta^{(I+J)\cdot(I+J)}\,.$$

For an ordered tuple $x = (x_1,\dots,x_n) \in \mathbb{Z}_d^n$, we define $f(x) = \zeta^{x_1^2+\cdots+x_n^2}$. To complete the proof, it suffices to show that $f : \mathbb{Z}_d^n \to \mathbb{C}$ is well defined. To do this, we will show that ζ^{x^2} depends only on the residue of $x \mod d$. If d is odd, we can choose ζ to be a d^{th} root of unity and the claim is clear. If d is even $\zeta^{(x+d)^2} = \zeta^{x^2}\zeta^{2dx+d^2}$ but since $2dx+d^2$ is a multiple of $2d$, we get that $\zeta^{2dx+d^2} = 1$ and thus $\zeta^{(x+d)^2} = \zeta^{x^2}$. □

3 Non-rigidity of generalized Walsh–Hadamard matrices

In this section, we show that the GWH matrix $H_{d,n}$ becomes highly non-rigid for large values of n. The precise result is stated below.

Theorem 3.1. *Let $N = d^n$ for positive integers d,n. Let $0 < \varepsilon < 0.01$ and assume $n \geq 1/\psi$ where*

$$\psi = \frac{\varepsilon^2}{400\log^2(1/\varepsilon)d\log d}\,.$$

Then

$$r_{H_{d,n}}\left(N^{1-\psi}\right) \leq N^{\varepsilon}\,.$$

First we prove a few lemmas about symmetric polynomials that we will use in the proof of Theorem 3.1.

Lemma 3.2. *Let T_m denote the set of ordered tuples in $\mathbb{Z}_d^n$ such that at least m entries are equal to 0. Let $\mathsf{rep}(T_m) = \{I_1,\dots,I_k\}$. Consider the polynomials $P_1(x_1,\dots,x_n),\dots,P_k(x_1,\dots,x_n)$ defined by*

$$P_i(x_1,\dots,x_n) = \sum_{I\in\mathsf{perm}(I_i)} x^I\,.$$

For any complex numbers $y_1,\dots,y_m$, and any polynomial $Q(x_{m+1},\dots x_n)$ that is symmetric and degree at most $d-1$ in each of its variables, there exist coefficients $c_1,\dots,c_k$ such that

$$Q(x_{m+1},\dots,x_n) = \sum c_i P_i(y_1,\dots y_m,x_{m+1},\dots,x_n)\,.$$

Proof. It suffices to prove the statement for all Q of the form

$$\sum_{I''\in\mathsf{perm}(I')} x^{I''}$$

where $I' \in \mathbb{Z}_d^{n-m}$. We will prove this by induction on the degree. Clearly one of the I_i is $(0,0\dots0)$, so one of the polynomials $P_i(x_1,\dots,x_n)$ is constant. This finishes the case when Q has degree 0. Now we do the induction step. Note that we can extend I' to an element of T_m by setting the first m entries equal to 0. Call this extension I and say that $I \in \mathsf{perm}(I_i)$. We have

$$\sum_{I'' \in \mathsf{perm}(I')} x^{I''} = P_i(y_1, \ldots, y_m, x_{m+1}, \ldots, x_n) - R(y_1, \ldots, y_m, x_{m+1}, \ldots x_n)\,.$$

$R(y_1, \ldots, y_m, x_{m+1}, \ldots x_n)$, when viewed as a polynomial in $x_{m+1}, \ldots, x_n$ (since $y_1, \ldots, y_m$ are complex numbers that we can plug in), is symmetric and of lower degree than the left hand side. Thus, using the induction hypothesis, we can write R in the desired form. This completes the induction step. □

The key ingredient in the proof of Theorem 3.1 is the following lemma which closely resembles the main result in [7], but deals with matrices over $\mathbb{C}$.

Lemma 3.3. *Let $f : \mathbb{Z}_d^n \to \mathbb{C}$ be a symmetric function on the n variables. Let $N = d^n$. Let $0 < \varepsilon < 0.01$ and assume $n \geq 1/\psi$ where*

$$\psi = \frac{\varepsilon^2}{400 \log^2(1/\varepsilon) d \log d}\,.$$

Then

$$r_{M(f)}\left(N^{1-\psi}\right) \leq N^{\varepsilon}\,.$$

Let

$$\delta = \frac{\varepsilon}{10 \log(1/\varepsilon)}, \quad \text{and} \quad m = \left\lceil n\left(\frac{1-\delta}{d}\right)\right\rceil$$

and let S denote the set of all ordered tuples $(i_1, i_2, \ldots, i_n) \in \mathbb{Z}_d^n$ such that the entries indexed $1, 2, \ldots, m$ are equal to 0, the entries indexed $m+1, \ldots, 2m$ are equal to 1 and in general for $0 \leq i \leq d-1$, the entries indexed $im+1, \ldots, (i+1)m$ are equal to i. Note $|S| = d^{n-dm} \approx d^{\delta n} = N^{\varepsilon^2}$ (since $n - dm$ is approximately δn).

The main idea will be to change f in a small number of locations so that it has many zeros in the set $\{\omega^{[I]} \mid I \in \mathbb{Z}_d^n\}$ in order to make use of Claim 2.20. More precisely, first we will change f to f' by changing its values in at most N^{ε} places so that f' is still symmetric in all of the variables and

$$\forall I \in S, \qquad P_{f'}\left(\omega^{[I]}\right) = 0\,.$$

Note that although the size of S is small, the fact that f' is symmetric implies that f' also vanishes on $\mathsf{perm}(S)$, which covers almost all of $\mathbb{Z}_d^n$. Once we have shown the above, we quantitatively bound the number of entries changed between $M(f)$ and $M(f')$ and also the rank of $M(f')$ to complete the proof of Lemma 3.3. To do the first part, we need the following sub-lemma.

Lemma 3.4. *Let T denote the set of all ordered tuples $(i_1, i_2, \ldots, i_n) \in \mathbb{Z}_d^n$ such that at least $n\big(1-\delta\big)$ of the entries are 0. By changing the values of f only on elements of T, we can obtain f' satisfying*

$$\forall I \in S, \qquad P_{f'}\left(\omega^{[I]}\right) = 0\,. \tag{3.1}$$

Proof. We interpret (3.1) as a system of linear equations where the unknowns are the values of f' at various points. Let $\mathsf{rep}(T) = \{J_1, J_2, \ldots, J_k\}$ for $J_1, J_2, \ldots J_k \in T$. Since we must maintain that f' is symmetric, there are essentially k variables each corresponding to an equivalence class of ordered tuples under permutations. Each equivalence class is of the form $\mathsf{perm}(J_j)$ and we denote the corresponding variable by m_j. The system of equations in (3.1) can be rewritten in the form

$$\forall I \in S \qquad \sum_{j=1}^{k} m_j \sum_{J \in \mathsf{perm}(J_j)} \omega^{I \cdot J} + \sum_{J' \notin T} f(J') \omega^{I \cdot J'} = 0 \,.$$

If we let $\mathsf{rep}(S) = \{I_1, I_2, \ldots, I_l\}$, the system has exactly l distinct equations corresponding to each element of $\mathsf{rep}(S)$ due to our symmetry assumptions. Let M denote the $l \times k$ coefficient matrix represented by $M_{ij} = \sum_{J \in \mathsf{perm}(J_j)} \omega^{I_i \cdot J}$. To show that the system has a solution, it suffices to show that the column span of M is full. This is equivalent to showing that for each $i = 1, 2, \ldots, l$ there exist coefficients $a_1, a_2, \ldots, a_k$ such that

$$\sum_{j=1}^{k} a_j \cdot \sum_{J \in \mathsf{perm}(J_j)} \omega^{I_i \cdot J} \neq 0 \,,$$

$$\forall i' \neq i \qquad \sum_{j=1}^{k} a_j \cdot \sum_{J \in \mathsf{perm}(J_j)} \omega^{I_{i'} \cdot J} = 0 \,.$$

Fix an index i_0. We can view each equation above as a polynomial in $\omega^{[I_i]}$ given by

$$P(x_1, \ldots, x_n) = \sum_{j=1}^{k} a_j \sum_{J \in \mathsf{perm}(J_j)} x^J$$

and the problem becomes equivalent to constructing a polynomial that vanishes on $\omega^{[I_i]}$ if and only if $i \neq i_0$. Note that only the entries $x_{dm+1}, \ldots, x_n$ matter as we have

$$x_1 = \cdots = x_m = 1, \ldots, x_{(d-1)m+1} = \cdots = x_{dm} = \omega^{d-1}$$

for all points we consider.

For $I_i = (i_1, i_2, \ldots i_n)$, let I'_i denote the (ordered) sub-tuple $(i_{dm+1}, \ldots, i_n)$. The problem is equivalent to constructing a polynomial

$$Q(x_{dm+1}, \ldots, x_n) = P(1, 1, \ldots, \omega^{d-1}, \ldots, \omega^{d-1}, x_{dm+1}, \ldots x_n)$$

such that Q vanishes on $\omega^{[I'_i]}$ if and only if $i \neq i_0$.

Lemma 3.2 implies that by choosing the coefficients $a_1, \ldots, a_k$, we can make Q be any polynomial that is symmetric in $x_{dm+1}, \ldots, x_n$ and degree at most $d-1$ in each of the variables.

Now consider the polynomial

$$Q_{i_0}(x_{dm+1}, \ldots, x_n) = \sum_{I' \in \mathsf{perm}(I'_{i_0})} \left(\frac{x_{dm+1}^d - 1}{x_{dm+1} - \omega^{I'(1)}} \right) \cdots \left(\frac{x_n^d - 1}{x_n - \omega^{I'(n-dm)}} \right).$$

Note this is a polynomial with coefficients in $\mathbb{C}$ since each of the factors reduces to a degree $d-1$ polynomial.

It is clear that the above polynomial is symmetric in all of the variables and satisfies the degree constraint so we know we can choose suitable coefficients $a_1,\dots,a_k$. We claim that the polynomial we construct does not vanish on $\omega^{[I'_{i_0}]}$ but vanishes on $\omega^{[I'_i]}$ for $i \neq i_0$. Indeed, the product

$$\left(\frac{x_{dm+1}^d - 1}{x_{dm+1} - \omega^{I'^{(1)}}}\right)\cdots\left(\frac{x_n^d - 1}{x_n - \omega^{I'^{(n-dm)}}}\right)$$

is 0 if and only if $(x_{dm+1},\dots,x_n) \neq I'$. However, there is exactly one $I' \in \mathsf{perm}(I'_{i_0})$ with $I' = I'_{i_0}$ and none with $I' = I'_i$ for $i \neq i_0$ since $I_1, I_2, \dots, I_l$ are representatives of distinct equivalence classes under permutation of entries. This means that the polynomial Q_{i_0} we constructed has the desired properties and completes the proof that the system is solvable. □

Proof of Lemma 3.3. Since $M(f) = (M(f) - M(f')) + M(f')$, to complete the proof of Lemma 3.3, it suffices to bound the number of nonzero entries in $M(f) - M(f')$ and the rank of $M(f')$.

The number of nonzero entries in each row and column of $(M(f) - M(f'))$ is at most $|T|$. This is exactly the number of elements of $\mathbb{Z}_d^n$ with at least $n\big(1-\delta\big)$ entries equal to 0. Using standard tail bounds on the binomial distribution (see [3]), the probability of a random ordered n-tuple having at least that many 0s is at most

$$\begin{aligned}\exp\left(-nD\left(1-\delta \mid\mid \frac{1}{d}\right)\right) &= \exp\left(-n\left((1-\delta)\log(d(1-\delta)) + \delta\log\left(\frac{d\delta}{d-1}\right)\right)\right)\\ &= d^{-n(1-\delta)}\exp\left(-n\left((1-\delta)\log(1-\delta) + \delta\log\left(\frac{d\delta}{d-1}\right)\right)\right)\end{aligned}$$

where

$$D(a \mid\mid b) = a\log\frac{a}{b} + (1-a)\log\frac{1-a}{1-b}$$

denotes the KL-divergence between Bernoulli distributions with means a and b.

For $\delta < 0.01$, the above is at most $d^{-n(1-4\delta\log(1/\delta))}$. Since $4\delta\log(1/\delta) < \varepsilon$, we change at most $d^{\varepsilon n}$ entries in each row and column.

By Claim 2.20, the rank of $M(f')$ is at most $d^n - |\mathsf{perm}(S)|$. Equivalently, this is the number of ordered n-tuples such that some element in $\{0,1,\dots,d-1\}$ appears less than $\frac{(1-\delta)n}{d}$ times. We use the multiplicative Chernoff bound and then union bound over the d possibilities to get the probability that a randomly chosen ordered n-tuple in $\mathbb{Z}_d^n$ is outside $\mathsf{perm}(S)$ is at most

$$d\exp\left(-\frac{\delta^2 n}{2d}\right) = \exp\left(-\frac{\delta^2 n}{2d} + \log d\right).$$

When $n > \frac{4d(\log d)}{\delta^2}$, the above is at most $d^{-(\delta^2 n)/(4d\log d)}$ and thus the rank of $M(f')$ is at most $d^{(1-\psi)n}$ where

$$\psi = \frac{\varepsilon^2}{400\log^2(1/\varepsilon)d\log d},$$

completing the proof of Lemma 3.3. □

Proof of Theorem 3.1. Combine Claim 2.22 and Lemma 3.3. □

Using Theorem 3.1, Lemma 2.21, and Claim 2.18, we get the following result which extends Lemma 3.3 to matrices where f is not symmetric.

Corollary 3.5. *For any function $f : \mathbb{Z}_d^n \to \mathbb{C}$ and any $0 < \varepsilon < 0.01$ such that $n \geq 1/\psi$ where*

$$\psi = \frac{\varepsilon^2}{400 \log^2(1/\varepsilon) d \log d},$$

we have

$$r_{M(f)}\left(2N^{1-\psi}\right) \leq N^{2\varepsilon}$$

where $N = d^n$.

4 Non-rigidity of DFT matrices of well-factorable size

Our goal in this section is to show that we can find infinitely many values of N for which the DFT matrix DFT_N is highly non-rigid. The integers N we analyze will be products of many distinct primes q_i with the property that $q_i - 1$ is smooth (has all prime factors small). For these values of N, we can decompose the matrix DFT_N into several submatrices that are closely related to Hadamard matrices. We then apply the results from the previous section to show that each submatrix is non-rigid and aggregate over the submatrices to conclude that DFT_N is non-rigid.

We first show precisely how to construct N. We rely on the following number theoretic result, found in [5], that allows us to find a large set of primes q_i for which $q_i - 1$ is smooth.

Definition 4.1. For a positive integer m, let $\rho^+(m)$ denote the largest prime factor of m. For a fixed positive integer a, let

$$\pi_a(x,y) = |\{p \mid a < p \leq x, \rho^+(p-a) \leq y\}|$$

where p ranges over all primes. In other words, $\pi_a(x,y)$ is the number of primes at most x such that $p-a$ is y-smooth.

Theorem 4.2 ([5])**.** *There exist constants x_0, C such that for $\beta = 0.2961, x > x_0$ and $y \geq x^\beta$ we have* [5]

$$\pi_1(x,y) > \frac{x}{(\log x)^C}.$$

Throughout the remainder of this paper, set $C_0 = C + 1$ where C is the constant in Theorem 4.2. The properties that we want N to have are stated in the following two definitions.

Definition 4.3. We say a prime q is (α, x)-good if the following conditions hold.

- $\frac{x}{(\log x)^{C_0}} \leq q \leq x$.
- All prime powers dividing $q-1$ are at most x^α.

[5] [5] proves the same inequality with $\pi_a(x,y)$ for any integer a where x_0 may depend on a and C is an absolute constant.

Definition 4.4. We say an integer N is (l,α,x)-factorable if the following conditions hold.

- $N = q_1 \cdots q_l$ where $q_1,\dots,q_l$ are distinct primes.
- $q_1,\dots,q_l$ are all (α,x)-good.

To show the existence of (l,α,x)-factorable integers, it suffices to show that there are many (α,x)-good primes. This is captured in the following lemma.

Lemma 4.5. *For a fixed constant C_0, any parameter $\alpha > 0.2961$, and sufficiently large x (possibly depending on α), there are at least $10x/(\log x)^{C_0}$ distinct (α,x)-good primes.*

Proof of Lemma 4.5. Let $y = x^\beta$ where $\beta = 0.2961$. By Theorem 4.2, for sufficiently large x, we can find at least
$$\left\lceil \frac{x}{(\log x)^C} - \frac{x}{(\log x)^{C_0}} \right\rceil$$
primes $p_1,\dots,p_l$ between $x/(\log x)^{C_0}$ and x such that all prime factors of $p_i - 1$ are at most x^β. Eliminate all of the p_i such that one of the prime powers in the prime factorization of $p_i - 1$ is more than x^α. Note that there are at most $x^\beta \log x$ integers in the range $[x^\alpha, x]$ that are powers of primes smaller than x^β. Each of these prime powers can divide at most $x^{1-\alpha}$ of the elements $\{p_1 - 1,\dots,p_l - 1\}$, so in total, we eliminate at most $x^{1-\alpha+\beta}\log x$ of the p_i. Thus, for sufficiently large x, the number of (α,x)-good primes is at least
$$\frac{x}{(\log x)^C} - \frac{x}{(\log x)^{C_0}} - x^{1-\alpha+\beta}\log x \geq \frac{x}{2(\log x)^C}. \qquad \square$$

For simplicity, we will set $\alpha = 0.3$ by default.

Definition 4.6. We say a prime is x-good if it is $(0.3,x)$-good. We say an integer N is (l,x)-factorable if it is $(l,0.3,x)$-factorable.

Lemma 4.5 implies that for all sufficiently large x and $l \leq \frac{x}{(\log x)^{C_0}}$ (where C_0 is an absolute constant), we can find (l,x)-factorable integers. We now show that if we choose x sufficiently large and N to be (l,x)-factorable for some
$$\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}},$$
then DFT_N is highly non-rigid.

Theorem 4.7. *Let $0 < \varepsilon < 0.01$ be some constant. For x sufficiently large and N a (l,x)-factorable number with*
$$\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}},$$
we must have
$$r_{\mathrm{DFT}_N}\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.36}\right)}\right) \leq N^{7\varepsilon}.$$

In order to prove Theorem 4.7, we will first prove a series of preliminary results that characterize the structure of DFT and GWH matrices.

4.1 Structure of generalized Walsh–Hadamard and DFT matrices

Lemma 4.8. *Let $n = x_1 x_2 \cdots x_j$ for pairwise relatively prime positive integers $x_1, \ldots, x_j$. There exists a permutation of the rows and columns of DFT_n, say DFT', such that*

$$\mathrm{DFT}' = \mathrm{DFT}_{x_1} \otimes \cdots \otimes \mathrm{DFT}_{x_j}\,,$$

where $\otimes$ denotes the Kronecker product.

Proof. This follows from Fact 2.11. $\square$

Lemma 4.9. *Let $M = A \otimes B$ where A is an $m \times m$ matrix and B is an $n \times n$ matrix. For any two integers r_1, r_2 we have*

$$\mathsf{r}_M(r_1 n + r_2 m) \le \mathsf{r}_A(r_1)\mathsf{r}_B(r_2)\,.$$

Proof. The proof of this lemma is similar to the proof of Lemma 2.21. There are matrices E, F with at most $\mathsf{r}_A(r_1)$ and $\mathsf{r}_B(r_2)$ nonzero entries respectively such that $\mathsf{rank}(A+E) \le r_1$ and $\mathsf{rank}(B+F) \le r_2$. We will now show that $\mathsf{rank}(M - E \otimes F) \le r_1 n + r_2 m$. Indeed

$$M - E \otimes F = (A+E) \otimes B - E \otimes (B+F)$$

and the right hand side of the above has rank at most $r_1 n + r_2 m$ since rank multiplies under the Kronecker product. Clearly $E \otimes F$ has at most $\mathsf{r}_A(r_1)\mathsf{r}_B(r_2)$ nonzero entries in each row and column so we are done. $\square$

Lemma 4.10. *Consider the matrix*

$$A = (\underbrace{\mathrm{DFT}_{t_1} \otimes \cdots \otimes \mathrm{DFT}_{t_1}}_{a_1}) \otimes \cdots \otimes (\underbrace{\mathrm{DFT}_{t_n} \otimes \cdots \otimes \mathrm{DFT}_{t_n}}_{a_n})\,.$$

Let $0 < \varepsilon < 0.01$ be some chosen parameter and D be some sufficiently large constant (possibly depending on ε). Assume $t_1 \le t_2 \le \cdots \le t_n$ and $a_i \ge \max\left(\frac{t_i^2(\log t_i)^2}{\varepsilon^{10}}, D\right)$ for all i. Let $P = t_1^{a_1} \cdots t_n^{a_n}$ and $L = \lceil 2 \log\log P \rceil$. Then

$$\mathsf{r}_A\left(P^{1-\varepsilon^6/(10Lt_n^2 \log t_n)}\right) \le P^{5\varepsilon}\,.$$

Proof. First, we consider the case when there exists an integer B such that $B \le t_1^{a_1}, \ldots, t_n^{a_n} \le B^2$. Note that

$$(\underbrace{\mathrm{DFT}_{t_i} \otimes \cdots \otimes \mathrm{DFT}_{t_i}}_{a_i}) = H_{t_i, a_i}\,.$$

By Theorem 3.1, for each i there exists a matrix E_i such that E_i has at most $t_i^{\varepsilon a_i}$ nonzero entries in each row and column and

$$\mathsf{rank}(H_{t_i,a_i} - E_i) \le t_i^{a_i\left(1-\varepsilon^4/(t_i^2 \log t_i)\right)}\,.$$

Let $A_i = H_{t_i,a_i} - E_i$. Then

$$(\underbrace{\mathrm{DFT}_{t_1} \otimes \cdots \otimes \mathrm{DFT}_{t_1}}_{a_1}) \otimes \cdots \otimes (\underbrace{\mathrm{DFT}_{t_n} \otimes \cdots \otimes \mathrm{DFT}_{t_n}}_{a_n}) = (E_1 + A_1) \otimes \cdots \otimes (E_n + A_n)$$

$$= \sum_{S \subset [n]} \left(\bigotimes_{i \in S} A_i \right) \otimes \left(\bigotimes_{i' \notin S} E_{i'} \right) = \sum_{S \subset [n], |S| \geq \varepsilon n} \left(\bigotimes_{i \in S} A_i \right) \otimes \left(\bigotimes_{i' \notin S} E_{i'} \right) + \sum_{S \subset [n], |S| < \varepsilon n} \left(\bigotimes_{i \in S} A_i \right) \otimes \left(\bigotimes_{i' \notin S} E_{i'} \right).$$

Let the first term above be N_1 and the second term be N_2. We bound the rank of N_1 and the number of nonzero entries in each row and column of N_2. Note that by grouping terms in the sum for N_1, we can find matrices E_S for all $S \subset [n]$ with $|S| = \varepsilon n$ and write

$$N_1 = \sum_{S \subset [n], |S| = \varepsilon n} \left(\bigotimes_{i \in S} A_i \right) \otimes E_S .$$

Now we have

$$\begin{aligned} \mathsf{rank}(N_1) &\leq \sum_{S \subset [n], |S| = \varepsilon n} P \prod_{i \in S} \frac{1}{t_i^{a_i \varepsilon^4/(t_i^2 \log t_i)}} \leq \binom{n}{\varepsilon n} \frac{P}{\left(B^{\varepsilon^4/(t_n^2 \log t_n)}\right)^{\varepsilon n}} \\ &\leq \frac{(n)^{\varepsilon n}}{\left(\frac{\varepsilon n}{3}\right)^{\varepsilon n}} \frac{P}{\left(B^{\varepsilon^4/(t_n^2 \log t_n)}\right)^{\varepsilon n}} \leq \left(\frac{3}{\varepsilon}\right)^{\varepsilon n} \frac{P}{B^{\varepsilon^5 n/(t_n^2 \log t_n)}} . \end{aligned}$$

Since we assumed $a_i \geq \max\left(\frac{t_i^2 (\log t_i)^2}{\varepsilon^{10}}, D\right)$, we get

$$B^{\varepsilon^4/(t_n^2 \log t_n)} \geq t_n^{a_n \varepsilon^4/(2t_n^2 \log t_n)} \geq \max\left(t_n^{0.5 \log t_n}, t_n^{D\varepsilon^4/(2t_n^2 \log t_n)}\right) .$$

Either the first term is larger than $(3/\varepsilon)^2$ or t_n is bounded above by some function of ε in which case if we choose D sufficiently large, the second term will be larger than $(3/\varepsilon)^2$. In any case we get

$$\mathsf{rank}(N_1) \leq \frac{P}{\left(\frac{\varepsilon}{3} \cdot B^{\varepsilon^4/(t_n^2 \log t_n)}\right)^{\varepsilon n}} \leq \frac{P}{B^{\varepsilon^5 n/(2t_n^2 \log t_n)}} \leq P^{1-\varepsilon^5/(4t_n^2 \log t_n)} .$$

Now we bound the number of nonzero entries in each row and column of N_2. This number is at most

$$2^n B^{2\varepsilon n} P^{\varepsilon} \leq 2^n P^{3\varepsilon} \leq P^{4\varepsilon} .$$

Thus, when we have $B \leq t_1^{a_1}, \ldots, t_n^{a_n} \leq B^2$,

$$\mathsf{r}_A\left(P^{1-\varepsilon^5/(4t_n^2 \log t_n)}\right) \leq P^{4\varepsilon} .$$

Now we move on to the case where we no longer have control over the range of values $t_1^{a_1}, \ldots, t_n^{a_n}$. Fix $k = 2^D$ and consider the intervals $I_1 = [k, k^2), I_2 = [k^2, k^4), \ldots, I_j = [k^{2^{j-1}}, k^{2^j}), \ldots$ and so on. Note

$$A = \bigotimes_{i \in [L]} \left(\bigotimes_{t_j^{a_j} \in I_i} (\underbrace{\mathrm{DFT}_{t_j} \otimes \cdots \otimes \mathrm{DFT}_{t_j}}_{a_j}) \right) .$$

For an integer i, let $P_i = \prod_{t_j^{a_j} \in I_i} t_j^{a_j}$. Let T be the set of indices $i \in [L]$ such that $P_i \geq P^{\varepsilon/(2L)}$. Then

$$A = \left(\bigotimes_{i \in T} \left(\bigotimes_{t_j^{a_j} \in I_i} \underbrace{(\mathrm{DFT}_{t_j} \otimes \cdots \otimes \mathrm{DFT}_{t_j})}_{a_j} \right) \right) \otimes \left(\bigotimes_{i \notin T} \left(\bigotimes_{t_j^{a_j} \in I_i} \underbrace{(\mathrm{DFT}_{t_j} \otimes \cdots \otimes \mathrm{DFT}_{t_j})}_{a_j} \right) \right) = B \otimes C$$

where naturally B denotes the first term and C denotes the second.

Note that the dimension of the matrix C, which we denote by $|C|$, is at most $\left(P^{\varepsilon/(2L)}\right)^L = P^{\varepsilon/2}$. We now apply Lemma 4.9 repeatedly to bound the rigidity of B. Let

$$B_i = \left(\bigotimes_{t_j^{a_j} \in I_i} \underbrace{(\mathrm{DFT}_{t_j} \otimes \cdots \otimes \mathrm{DFT}_{t_j})}_{a_j} \right).$$

Then we have,

$$\mathsf{r}_B \left(\left(\prod_{i \in T} P_i \right) \left(\sum_{i \in T} \frac{1}{P_i^{\varepsilon^5/(4t_n^2 \log t_n)}} \right) \right) \leq \left(\prod_{i \in T} P_i \right)^{4\varepsilon}.$$

From the above inequality, the fact that $P_i \geq P^{\varepsilon/(2L)}$ for all $i \in T$, and $|C| \leq P^{\varepsilon/2}$ we deduce

$$\mathsf{r}_A \left(P^{1-\varepsilon^6/(10Lt_n^2 \log t_n)} \right) \leq \mathsf{r}_A \left(LP^{1-\varepsilon^6/(8Lt_n^2 \log t_n)} \right) \leq |C| \mathsf{r}_B \left(\frac{LP^{1-\varepsilon^6/(8Lt_n^2 \log t_n)}}{|C|} \right) \leq P^{5\varepsilon}. \qquad \square$$

4.2 Proof of Theorem 4.7

To complete the proof of Theorem 4.7, we will break DFT_N into submatrices, show that each submatrix is non-rigid using techniques from the previous section, and then combine our estimates to conclude that DFT_N is non-rigid. Recall that N is (l,x)-factorable with

$$\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}},$$

meaning $N = q_1 q_2 \cdots q_l$ for some distinct primes $q_1, \ldots, q_l$ where $q_i - 1$ has no large prime power divisors for all i. Let γ be a primitive N^{th} root of unity.

Definition 4.11. For a subset $S \subset [l]$ define $\mathsf{mult}_N(S) = \prod_{s \in S} q_s$ and $\mathsf{fact}_N(S) = \prod_{s \in S}(q_s - 1)$.

Definition 4.12. For all $S \subset [l]$ we will define T_S as the subset of $[N] \times [N]$ indexed by (i,j) such that

$$\begin{aligned} &\forall s \in S \qquad ij \not\equiv 0 \mod q_s\,, \\ &\forall s \notin S \qquad ij \equiv 0 \mod q_s\,. \end{aligned}$$

Note that as S ranges over all subsets of $[l]$, the sets T_S form a partition of $[N] \times [N]$.

For each S, we will divide the set T_S into submatrices such that when filled with the corresponding entries of DFT_N , we can apply Lemma 4.10 to show that each submatrix is nonrigid. The key intuition is that for a given prime q_i, once we restrict to nonzero residues, the multiplicative subgroup actually has the additive structure of $\mathbb{Z}_{q_i-1}$. Since $q_i - 1$ is smooth, $\mathbb{Z}_{q_i-1}$ is a direct sum of cyclic groups of small order.

Definition 4.13. For all $S \subset [l]$, we define the $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ matrix $M(S)$ as follows. Let R_S be the set of residues modulo $\mathsf{mult}_N(S)$ that are relatively prime to $\mathsf{mult}_N(S)$. Note that $|R_S| = \mathsf{fact}_N(S)$. Each row and each column of $M(S)$ is indexed by an element of R_S and the entry in row i and column j is $\theta^{i\cdot j}$ where θ is a primitive $\mathsf{mult}_N(S)$ root of unity. The exact order of the rows and columns will not matter for our uses. Note that replacing θ with θ^k for k relatively prime to $\mathsf{mult}_N(S)$ simply permutes the rows so it does not matter which root of unity we choose.

Lemma 4.14. *Consider the set of entries in* DFT_N *indexed by elements of* T_S*. We can partition this set into* $\prod_{s\notin S}(2q_s - 1)$ *submatrices each of size* $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ *that are equivalent to* $M(S)$ *up to some permutation of rows and columns.*

Proof. In T_S, for each prime q_s with $s \notin S$, there are $2q_s - 1$ choices for what i and j are $\mod q_s$. Now fix the choice of $i, j \mod q_s$ for all $s \notin S$. We restrict to indices with $i \equiv c_1 \mod \prod_{s\notin S} q_s$ and $j \equiv c_2 \mod \prod_{s\notin S} q_s$ for some c_1, c_2.

We are left with a $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ matrix, call it A, where i and j run over all residues modulo $\mathsf{mult}_N(S)$ that are relatively prime to $\mathsf{mult}_N(S)$. Naturally, label all rows and columns of this matrix by what the corresponding indices i and j are modulo $\mathsf{mult}_N(S)$. For a row labeled a and a column labeled b, we compute the entry A_{ab}. The value is $\gamma^{a'\cdot b'}$ where a' is the unique element of $\mathbb{Z}_N$ such that $a' \equiv a \mod \mathsf{mult}_N(S)$ and $a' \equiv c_1 \mod \prod_{s\notin S} q_s$ and b' is defined similarly. We have

$$a' \cdot b' \equiv ab \mod \mathsf{mult}_N(S),$$
$$a' \cdot b' \equiv c_1c_2 \equiv 0 \mod \prod_{s\notin S} q_s .$$

Therefore

$$a'b' \equiv k \prod_{s\notin S} q_s ab \mod \mathsf{mult}_N(S)$$

where k is defined as an integer such that $k\prod_{s\notin S} q_s \equiv 1 \mod \mathsf{mult}_N(S)$. Note that k clearly exists since $\prod_{s\notin S} q_s$ and $\mathsf{mult}_N(S)$ are relatively prime. Since $\gamma^{k\prod_{s\notin S} q_s}$ is a primitive $\mathsf{mult}_N(S)$ root of unity, the matrix A is equivalent to $M(S)$ up to some permutation, as desired. $\square$

Lemma 4.15. *For a subset* $S \subset [l]$ *with* $|S| = k$ *and* $M(S)$ *(as defined in Definition 4.13) a* $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ *matrix as described above. we have*

$$r_{M(S)}\left(\frac{\mathsf{fact}_N(S)}{\exp(\varepsilon^6 x^{0.37})}\right) \le (\mathsf{fact}_N(S))^{6\varepsilon}$$

as long as $k \ge \frac{x}{(\log x)^{C_0+200}}$.

Proof. Without loss of generality $S = \{1, 2, \dots, k\}$. Consider the factorizations of $q_1 - 1, \dots, q_k - 1$ into prime powers. For each prime power $p_i^{e_i} \le x^{0.3}$, let $c(p_i^{e_i})$ be the number of indices j for which $p_i^{e_i}$ appears (exactly) in the factorization of $q_j - 1$. Consider all prime powers $p_i^{e_i}$ for which $c(p_i^{e_i}) < x^{0.62}$.

$$\prod_{t, c(t) \le x^{0.62}} t^{c(t)} \le \left((x^{0.3})^{x^{0.62}}\right)^{x^{0.3}} \le x^{x^{0.92}} .$$

Now consider all prime powers say $t_1, \dots, t_n$ for which $c(t_i) \ge x^{0.62}$. Let $P = t_1^{c(t_1)} \cdots t_n^{c(t_n)}$. From the above we know that as long as x is sufficiently large

$$P \ge \frac{\mathsf{fact}_N(S)}{x^{x^{0.92}}} \ge (\mathsf{fact}_N(S))^{(1-\varepsilon)} \frac{\left(\frac{x}{(\log x)^{C_0+1}}\right)^{\varepsilon k}}{x^{x^{0.92}}} \ge (\mathsf{fact}_N(S))^{(1-\varepsilon)} . \tag{4.1}$$

We will use the prime powers t_i and Theorem 3.1 to show that $M(S)$ is not rigid. Note that we can associate each row and column of $M(S)$ with an ordered k-tuple $(a_1, \dots, a_k)$ where $a_i \in \mathbb{Z}_{q_i - 1}$ as follows. First, it is clear that each row and column of $M(S)$ can be associated with an ordered k-tuple $(z_1, \dots, z_k) \in \mathbb{F}_{q_1}^\times \times \cdots \times \mathbb{F}_{q_k}^\times$. Now $\mathbb{Z}_{q_i}^\times$ can be viewed as a cyclic group on $q_i - 1$ elements. This allows us to create a bijection between the rows and columns of $M(S)$ and elements of $\mathbb{Z}_{q_1 - 1} \times \cdots \times \mathbb{Z}_{q_k - 1}$.

Also note that for a row indexed by $A = (a_1, \dots, a_k)$ and a column indexed by $B = (b_1, \dots, b_k)$, the entry $M(S)_{AB}$ is dependent only on $A + B$. We will now decompose $M(S)$ into several $P \times P$ submatrices. In particular, we can write $q_i - 1 = d_i T_i$ where T_i is a product of some subset of $\{t_1, \dots, t_n\}$ and d_i is relatively prime to T_i. We have $T_1 T_2 \cdots T_k = P$. For each $A', B' \in \mathbb{Z}_{d_1} \times \cdots \times \mathbb{Z}_{d_k}$, we can construct a $P \times P$ submatrix $M(S, A', B')$ consisting of all entries $M(S)_{AB}$ of $M(S)$ such that $A \equiv A', B \equiv B'$ (where the equivalence is over $\mathbb{Z}_{d_1} \times \cdots \times \mathbb{Z}_{d_k}$). This gives us d^2 different submatrices where $d = d_1 \cdots d_k$. Naturally, we can associate each row and column of a submatrix $M(S, A', B')$ with an element of $\mathbb{Z}_{T_1} \times \cdots \times \mathbb{Z}_{T_k}$ such that for a row labeled I and a column labeled J, the entry $M(S, A', B')_{IJ}$ only depends on $I + J$. In particular, this means that $X\left(M(S, A', B')\right)X$ is diagonal where $X = \mathrm{DFT}_{T_1} \otimes \cdots \otimes \mathrm{DFT}_{T_k}$. Now, using Lemma 4.8, we can rewrite

$$X = \underbrace{(\mathrm{DFT}_{t_1} \otimes \cdots \otimes \mathrm{DFT}_{t_1})}_{c(t_1)} \otimes \cdots \otimes \underbrace{(\mathrm{DFT}_{t_n} \otimes \cdots \otimes \mathrm{DFT}_{t_n})}_{c(t_n)} .$$

Since for x sufficiently large, $c(t_i) \ge x^{0.62} \ge t_i^2 (\log t_i)^2 / \varepsilon^{10}$, we can use Lemma 4.10 and get that

$$\mathsf{r}_X\left(P^{1 - \varepsilon^6 / (20 (\log\log P) x^{0.62})}\right) \le P^{5\varepsilon} .$$

Let E be the matrix of changes to reduce the rank of X according to the above. We have that E has at most P^ε nonzero entries in each row and column, and

$$\mathsf{rank}(X - E) \le P^{1 - \varepsilon^6 / (20 (\log\log P) x^{0.62})} .$$

We can write $M(S)$ in block form as

$$\begin{bmatrix} M(S, A_1, B_1) & M(S, A_1, B_2) & \dots & M(S, A_1, B_d) \\ M(S, A_2, B_1) & M(S, A_2, B_2) & \dots & M(S, A_1, B_d) \\ \vdots & \vdots & \ddots & \vdots \\ M(S, A_d, B_1) & M(S, A_d, B_2) & \dots & M(S, A_d, B_d) \end{bmatrix}$$

where $A_1,\dots,A_d$ and $B_1,\dots,B_d$ range over the elements of $\mathbb{Z}_{d_1}\times\cdots\times\mathbb{Z}_{d_k}$. We can rearrange the above as

$$\begin{bmatrix} M(S,A_1,B_1) & \dots & M(S,A_1,B_d) \\ \vdots & \ddots & \vdots \\ M(S,A_d,B_1) & \dots & M(S,A_d,B_d) \end{bmatrix} = \begin{bmatrix} XD_{11}X & \dots & XD_{1d}X \\ \vdots & \ddots & \vdots \\ XD_{d1}X & \dots & XD_{dd}X \end{bmatrix}$$

where the D_{ij} are diagonal matrices. Now consider the matrix

$$E(S) = \begin{bmatrix} ED_{11}E & \dots & ED_{1d}E \\ \vdots & \ddots & \vdots \\ ED_{d1}E & \dots & ED_{dd}E \end{bmatrix}.$$

We have

$$M(S)-E(S) = \begin{bmatrix} XD_{11}X-ED_{11}E & \dots & XD_{1d}X-ED_{1d}E \\ \vdots & \ddots & \vdots \\ XD_{d1}X-ED_{d1}E & \dots & XD_{dd}X-ED_{dd}E \end{bmatrix} =$$
$$\begin{bmatrix} XD_{11}(X-E) & \dots & XD_{1d}(X-E) \\ \vdots & \ddots & \vdots \\ XD_{d1}(X-E) & \dots & XD_{dd}(X-E) \end{bmatrix} + \begin{bmatrix} (X-E)D_{11}E & \dots & (X-E)D_{1d}E \\ \vdots & \ddots & \vdots \\ (X-E)D_{d1}E & \dots & (X-E)D_{dd}E \end{bmatrix}.$$

In the above expression, each of the two terms has rank at most

$$dP^{1-\varepsilon^6/(20(\log\log P)x^{0.62})} = \frac{\mathsf{fact}_N(S)}{P^{\varepsilon^6/(20(\log\log P)x^{0.62})}} \le \frac{1}{2}\left(\frac{\mathsf{fact}_N(S)}{\exp\left(\varepsilon^6 x^{0.37}\right)}\right).$$

Note that when computing the rank, we only multiply by d (and not d^2) because the small blocks are all multiplied by the same low rank matrix on either the left or right. The number of nonzero entries in each row and column of $E(S)$ is at most $P^{5\varepsilon}d = \frac{\mathsf{fact}_N(S)}{P^{1-5\varepsilon}}$. Since $P \ge (\mathsf{fact}_N(S))^{1-\varepsilon}$, we conclude

$$\mathsf{r}_{M(S)}\left(\frac{\mathsf{fact}_N(S)}{\exp\left(\varepsilon^6 x^{0.37}\right)}\right) \le (\mathsf{fact}_N(S))^{6\varepsilon}. \qquad \square$$

We are now ready to complete the analysis of the non-rigidity of the DFT matrix DFT_N.

Proof of Theorem 4.7. Set the threshold $m = x^{0.365}$ and $k_0 = l-m$. The sets T_S, as S ranges over all subsets of $[l]$, form a partition of $[N]\times[N]$. For each $S\subset[l]$ with $|S|\ge k_0$, we will divide T_S into $\mathsf{fact}_N(S)\times\mathsf{fact}_N(S)$ submatrices using Lemma 4.14 and change entries to reduce the rank of every submatrix according to Lemma 4.15. We will not touch the entries in sets T_S for $|S|<k_0$. Call the resulting matrix M'. We now estimate the rank of M' and then the maximum number of entries changed in any row or column.

We remove all rows and columns corresponding to integers divisible by at least $\frac{m}{2}$ of the primes $q_1,\dots,q_l$. The number of rows and columns removed is at most

$$N\left(\sum_{S\subset[l],|S|=\frac{m}{2}}\prod_{i\in S}\frac{1}{q_i}\right) \le \frac{N}{\left(\frac{x}{(\log x)^{C_0}}\right)^{m/2}}\binom{l}{m/2} < N\left(\frac{l}{\frac{x}{(\log x)^{C_0}}}\right)^{m/2} \le \frac{N}{(\log x)^{x^{0.365}}}.$$

The remaining entries must be subdivided into matrices of the form $M(S)$ for various subsets $S \subset [l]$ with $|S| \geq k_0$. Let $q_1 < q_2 < \cdots < q_l$. The number of such submatrices is at most

$$\frac{N^2}{((q_1-1)\cdots(q_{k_0}-1))^2} \leq (q_{k_0+1}\cdots q_l)^2 \left(\frac{q_1\cdots q_{k_0}}{(q_1-1)\cdots(q_{k_0}-1)}\right)^2 \leq 3(q_{k_0+1}\cdots q_l)^2 \leq 3x^{2m}.$$

Each one of the submatrices has rank at most

$$\frac{N}{\exp\left(\varepsilon^6 x^{0.37}\right)},$$

so in total the rank is at most

$$N\frac{3x^{2m}}{\exp\left(\varepsilon^6 x^{0.37}\right)} \leq \frac{N}{\exp\left(\varepsilon^6 x^{0.369}\right)}.$$

Combining the two parts we easily get

$$\mathsf{rank}(M') \leq \frac{N}{\exp\left(\varepsilon^6 x^{0.365}\right)}.$$

Now we bound the number of entries changed. The number of entries changed in each row or column is at most

$$\frac{N}{((q_1-1)\cdots(q_{k_0}-1))}N^{6\varepsilon} \leq (q_{k_0+1}\cdots q_l)\left(\frac{q_1\cdots q_{k_0}}{(q_1-1)\cdots(q_{k_0}-1)}\right)N^{6\varepsilon} \leq 3N^{6\varepsilon+1.1m/l} \leq N^{7\varepsilon}.$$

As $\exp\left(\varepsilon^6 x^{0.365}\right) \geq \exp\left(\varepsilon^6 (\log N)^{0.36}\right)$ for sufficiently large x, we conclude

$$\mathsf{r}_{\mathrm{DFT}_N}\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.36}\right)}\right) \leq N^{7\varepsilon}. \qquad \square$$

5 Non-rigidity of all circulant matrices

In the previous section, we showed that there exists an infinite set of DFT matrices that are not Valiant-rigid. In this section, we will bootstrap the results from Section 4 to show that in fact, no (sufficiently large) DFT matrix is rigid.

The first ingredient will be a stronger form of Lemma 4.5. Recall that a prime q is defined to be x-good if $x/(\log x)^{C_0} \leq q \leq x$ and all prime powers dividing $q-1$ are at most $x^{0.3}$ and that an integer N is defined to be (l,x)-factorable if it can be written as the product of l distinct x-good primes.

To simplify our formulas we use the following notation.

Notation 5.1. Define

$$g_k(x) = \frac{x}{(\log x)^{C_0+k}}.$$

Lemma 5.2. *For all sufficiently large integers K, there exist l,x,N such that the following conditions hold:*

- *$g_{100}(x) \le l \le g_{10}(x)$,*
- *N is (l,x)-factorable,*
- *$K < N < K(\log K)^2$.*

Proof. Call an N well-factorable if it is (l,x)-factorable for some x and $g_{100}(x) \le l \le g_{10}(x)$. Let N_0 be the largest integer that is well-factorable with $N_0 \le K$. Assume N_0 is (l,x)-factorable.

We have $N_0 = q_1 \cdots q_l$ where $q_1, \dots, q_l$ are distinct, x-good primes. If $l < \lfloor g_{10}(x) \rfloor$ then by Lemma 4.5, we can find another x-good prime q_{l+1}. We can then replace N_0 with $q_{l+1}N_0$. $q_{l+1}N_0 > K$ by the maximality of N_0 and also $q_{l+1}N_0 \le N_0 x \le N_0(\log N_0)^2$ so $q_{l+1}N_0$ satisfies the desired conditions.

We now consider the case where $l = \lfloor g_{10}(x) \rfloor$. First, if $q_1, \dots, q_l$ are not the l largest x-good primes then we can replace one of them say q_1 with $q_1' > q_1$. The number $N' = q_1' q_2 \cdots q_l$ is well-factorable and between N_0 and $N_0(\log x)^{C_0}$. Using the maximality of N_0, we deduce that N' must be in the desired range.

On the other hand if $q_1, \dots, q_l$ are the l largest x-good primes, we know they are actually all between $3x/(\log x)^{C_0}$ and x. This is because by Lemma 4.5, there are at least $10x/(\log x)^{C_0}$ distinct x-good primes. Let $x' = 2x$. The above implies that $q_1, \dots, q_l$ are x'-good and clearly $g_{100}(x') \le l \le g_{10}(x')$. Furthermore, $g_{10}(x') > g_{10}(x) + 1$ so $l = \lfloor g_{10}(x) \rfloor < \lfloor g_{10}(x') \rfloor$ and we can now repeat the argument from the first case. □

We can now complete the proof that circulant matrices are not rigid.

Theorem 5.3. *Let $0 < \varepsilon < 0.01$ be a given parameter. For all sufficiently large N, if M is an $N \times N$ circulant (or Toeplitz) matrix, then*

$$\mathsf{r}_M\left(\frac{N}{\exp\left(\varepsilon^6 (\log N)^{0.35}\right)}\right) \le N^{15\varepsilon}.$$

Proof. First we analyze circulant matrices of size N_0 where N_0 is (l,x)-factorable for some $g_{100}(x) \le l \le g_{10}(x)$. Theorem 4.7 and Lemma 2.21 imply that for M_0 an $N_0 \times N_0$ circulant matrix where N_0 satisfies the previously mentioned conditions,

$$\mathsf{r}_{M_0}\left(\frac{2N_0}{\exp\left(\varepsilon^6 (\log N_0)^{0.36}\right)}\right) \le N_0^{14\varepsilon}.$$

Now for a circulant matrix M of arbitrary size $N \times N$, note that it is possible to embed an M in the upper left corner of a circulant matrix of any size at least $2N$. By Lemma 5.2, there exists an N_0 that is (l,x)-factorable for some $g_{100}(x) \le l \le g_{10}(x)$ such that

$$\frac{N_0}{(\log N_0)^2} \le N \le \frac{N_0}{2}.$$

We deduce

$$\mathsf{r}_M\left(\frac{2N_0}{\exp\left(\varepsilon^6 (\log N_0)^{0.36}\right)}\right) \le N_0^{14\varepsilon}.$$

Rewriting the bounds in terms of N we get

$$\mathsf{r}_M\left(\frac{N}{\exp\left(\varepsilon^6 (\log N)^{0.35}\right)}\right) \le N^{15\varepsilon}. \qquad \square$$

Remark 5.4. Note that our proof actually shows something slightly stronger, namely that the changes to reduce the rank of a circulant matrix are actually fixed linear combinations of the entries. See Definition 8.1 and Claim 8.2 for a more precise statement.

From the above and Claim 2.22, we immediately deduce that DFT matrices are not rigid.

Theorem 5.5. *Let $0 < \varepsilon < 0.01$ be a given parameter. For all sufficiently large N,*

$$\mathsf{r}_{\mathrm{DFT}_N}\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.35}\right)}\right) \le N^{15\varepsilon}\,.$$

6 Non-rigidity of G-circulant matrices for abelian groups

Using the results from the previous section, we can show that DFT_G and G-circulant matrices are not Valiant-rigid for any infinite class of finite abelian groups G. Our proof follows the same strategy as the proof of Lemma 4.10.

Theorem 6.1. *Let $0 < \varepsilon < 0.01$ be fixed. Let G be an abelian group and $f : G \to \mathbb{C}$ be a function. If $|G|$ is sufficiently large then*

$$\mathsf{r}_{\mathrm{DFT}_G}\left(\frac{|G|}{\exp\left(\varepsilon^8(\log|G|)^{0.32}\right)}\right) \le |G|^{19\varepsilon}\,.$$

Proof. By the Fundamental Theorem of Finite Abelian Groups we can write $G = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_a}$. By Fact 2.11, we have $\mathrm{DFT}_G = \mathrm{DFT}_{n_1} \otimes \cdots \otimes \mathrm{DFT}_{n_a}$. Let us write $F = \mathrm{DFT}_G$.

Without loss of generality, $n_1 \le n_2 \le \cdots \le n_a$. We will choose k to be a fixed, sufficiently large positive integer. By Theorem 5.5, we can ensure that for $N > k$

$$\mathsf{r}_{\mathrm{DFT}_N}\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.35}\right)}\right) \le N^{15\varepsilon}\,.$$

Consider the ranges $I_1 = [k,k^2), I_2 = [k^2,k^4), \ldots I_j = [k^{2^{j-1}},k^{2^j})\ldots$ and so on. Let S_j be a multiset defined by $S_j = I_j \cap \{n_1,\ldots,n_a\}$. Fix a j and let the elements of S_j be $x_1 \le \cdots \le x_b$. By Theorem 5.5, for each x_i, there are matrices E_{x_i} and A_{x_i} such that $\mathrm{DFT}_{x_i} = A_{x_i} + E_{x_i}$, E_{x_i} has at most $x_i^{15\varepsilon}$ nonzero entries in each row and column, and

$$\mathsf{rank}(A_{x_i}) \le \frac{x_i}{\exp\left(\varepsilon^6(\log x_i)^{0.35}\right)}\,.$$

Now we can write

$$M_j = \mathrm{DFT}_{x_1} \otimes \cdots \otimes \mathrm{DFT}_{x_b} = (A_{x_1} + E_{x_1}) \otimes \cdots \otimes (A_{x_b} + E_{x_b}) = \sum_{S\subset[b]}\left(\bigotimes_{i\in S} A_{x_i}\right) \otimes \left(\bigotimes_{i'\notin S} E_{x'_i}\right)$$

$$= \sum_{S\subset[b],|S|\ge\varepsilon b}\left(\bigotimes_{i\in S} A_{x_i}\right) \otimes \left(\bigotimes_{i'\notin S} E_{x'_i}\right) + \sum_{S\subset[b],|S|<\varepsilon b}\left(\bigotimes_{i\in S} A_{x_i}\right) \otimes \left(\bigotimes_{i'\notin S} E_{x'_i}\right).$$

Let the first term above be N_1 and the second term be N_2. We will bound the rank of N_1 and the number of nonzero entries in each row and column of N_2. Note that by grouping the terms in the sum for N_1 we can write it in the form

$$\sum_{S\subset[b],|S|=\lceil\varepsilon b\rceil} \bigotimes_{i\in S} A_{x_i}\otimes E_S$$

where E_S is some matrix for each S. This implies that

$$\begin{aligned}\mathsf{rank}(N_1) \le \binom{b}{\lceil\varepsilon b\rceil}\frac{x_1\cdots x_b}{(\exp(\varepsilon^6(\log x_1)^{0.35}))^{\lceil\varepsilon b\rceil}} &\le \frac{b^{\lceil\varepsilon b\rceil}}{(\frac{\varepsilon b}{3})^{\lceil\varepsilon b\rceil}}\frac{x_1\cdots x_b}{(\exp(\varepsilon^6(\log x_1)^{0.35}))^{\lceil\varepsilon b\rceil}}\\ &= x_1\cdots x_b\left(\frac{3}{\varepsilon\exp(\varepsilon^6(\log x_1)^{0.35})}\right)^{\lceil\varepsilon b\rceil}.\end{aligned}$$

As long as k is sufficiently large, we have

$$\begin{aligned}\mathsf{rank}(N_1) \le x_1\cdots x_b\left(\frac{3}{\varepsilon\exp(\varepsilon^6(\log x_1)^{0.35})}\right)^{\lceil\varepsilon b\rceil} &\le x_1\cdots x_b\left(\frac{1}{\exp(\varepsilon^6(\log x_1)^{0.34})}\right)^{\lceil\varepsilon b\rceil}\\ &\le \frac{x_1\cdots x_b}{\exp(\varepsilon^7(\log x_1\cdots x_b)^{0.33})}\end{aligned}$$

where in the last step we used the fact that $x_i \le x_1^2$ for all i. The number of nonzero entries in each row or column of N_2 is at most

$$2^b x_b\cdots x_{b-\lfloor\varepsilon b\rfloor+1}(x_{b-\lfloor\varepsilon b\rfloor}\cdots x_1)^{15\varepsilon} = 2^b(x_1\cdots x_b)^{15\varepsilon}(x_b\cdots x_{b-\lfloor\varepsilon b\rfloor+1})^{1-15\varepsilon} \le (x_1\cdots x_b)^{18\varepsilon}.$$

Note in the last step above, we used the fact that $x_i \le x_1^2$.

For each integer c between 2 and k, let m_c be the number of copies of c in the set $\{n_1,\dots,n_a\}$. If $m_c \ge k^2(\log k)^2/\varepsilon^4$ then by Theorem 3.1, if we define $A_c = \underbrace{\mathrm{DFT}_c\otimes\cdots\otimes\mathrm{DFT}_c}_{m_c}$ then

$$\mathsf{r}_{A_c}\left(c^{m_c(1-\varepsilon^4/(k^2\log k))}\right) \le c^{m_c\varepsilon}.$$

Let $L = \lceil 2\log\log|G|\rceil$ and ensure that $|G|$ is sufficiently large so that $L > k$. Let T be the set of integers c between 2 and k such that $c^{m_c} \ge |G|^{\varepsilon/(2L)}$ (note that as long as $|G|$ is sufficiently large, all elements of T must satisfy $m_c \ge k^2(\log k)^2/\varepsilon^4$). Let R be the set of indices j for which $\prod_{x\in S_j} x \ge |G|^{\varepsilon/(2L)}$. Since S_j is clearly empty for $j \ge L$, the matrix F can be written as

$$F = \left(\bigotimes_{2\le c<k}\left(\underbrace{\mathrm{DFT}_c\otimes\cdots\otimes\mathrm{DFT}_c}_{m_c}\right)\right)\otimes\left(\bigotimes_{1\le j\le L} M_j\right).$$

Define

$$B = \left(\bigotimes_{c\notin T}\left(\underbrace{\mathrm{DFT}_c\otimes\cdots\otimes\mathrm{DFT}_c}_{m_c}\right)\right)\otimes\left(\bigotimes_{j\notin R} M_j\right).$$

Note that the size of B is at most

$$\left(|G|^{\varepsilon/(2L)}\right)^{k+L} \le |G|^{\varepsilon}.$$

Also $F = B \otimes D$ where

$$D = \left(\bigotimes_{c\in T}\left(\underbrace{\mathrm{DFT}_c \otimes \cdots \otimes \mathrm{DFT}_c}_{m_c}\right)\right) \otimes \left(\bigotimes_{j\in R} M_j\right).$$

For any rank r, we have $\mathsf{r}_M(|B|r) \le |B|\mathsf{r}_D(r)$. Applying Lemma 4.9 iteratively, we get

$$\mathsf{r}_D\left(\frac{|G|}{|B|}\left(\sum_{c\in T}\frac{1}{c^{m_c\varepsilon^4/(k^2\log k)}} + \sum_{j\in R}\frac{1}{\exp\left(\varepsilon^7(\log\prod_{x\in S_j}x)^{0.33}\right)}\right)\right) \le \left(\frac{|G|}{|B|}\right)^{18\varepsilon}.$$

Note that

$$\left(\sum_{c\in T}\frac{1}{c^{m_c\varepsilon^4/(k^2\log k)}} + \sum_{j\in R}\frac{1}{\exp\left(\varepsilon^7(\log\prod_{x\in S_j}x)^{0.33}\right)}\right) \le \frac{k}{|G|^{\varepsilon^5/(2Lk^2\log k)}} + \frac{L}{\exp\left(\varepsilon^8(\log|G|/2L)^{0.33}\right)}$$
$$\le \frac{1}{\exp\left(\varepsilon^8(\log|G|)^{0.32}\right)}.$$

Overall, we conclude

$$\mathsf{r}_F\left(\frac{|G|}{\exp\left(\varepsilon^8(\log|G|)^{0.32}\right)}\right) \le |B|\left(\frac{|G|}{|B|}\right)^{18\varepsilon} \le |G|^{19\varepsilon}. \qquad \square$$

Theorem 6.2. *Let $0 < \varepsilon < 0.01$ be fixed. Let G be an abelian group and $f : G \to \mathbb{C}$ be a function. Let $M = M_G(f)$ be a G-circulant matrix. If $|G|$ is sufficiently large then*

$$r_M\left(\frac{2|G|}{\exp\left(\varepsilon^8(\log|G|)^{0.32}\right)}\right) \le |G|^{38\varepsilon}.$$

Proof. Note DFT_G diagonalizes M. Thus, combining Theorem 6.1 with Lemma 2.21 gives the desired conclusion. $\square$

The rigidity results we proved hold over $\mathbb{C}$. By examining the proofs more carefully, we can actually show that when G is an abelian group, the same results hold for G-circulant matrices over an abelian extension of $\mathbb{Q}$ of degree $\tilde{O}(N^3)$.

Theorem 6.3. *Let G be an abelian group of order N. Then there exists $m = \tilde{O}(N^3)$, depending only on G, such that G-circulant matrices with entries in $\mathbb{Q}$ satisfy*

$$r_M\left(\frac{2|G|}{\exp\left(\varepsilon^8(\log|G|)^{0.32}\right)}\right) \le |G|^{38\varepsilon}$$

over the m^{th} cyclotomic field. The same bound holds for the matrix DFT_G.

Proof. First note that it is immediate from our proof that the GWH matrix $H_{d,n}$ is not rigid over $\mathbb{Q}[\omega]$ where ω is a d^{th} primitive root of unity. Now for well-factorable integers $N = p_1 p_2 \cdots p_k$, the additional roots of unity that we need to adjoin for non-rigidity of DFT_N are all roots of unity with order dividing $(p_1 - 1)(p_2 - 1) \cdots (p_k - 1)$. Thus, DFT_N is not rigid over an extension $\mathbb{Q}[\omega][\alpha]$ where α is a root of unity of order at most N. Finally for other values of N, we find a well-factorable integer $N' = \tilde{O}(N)$ and embed DFT_N into an $N' \times N'$ circulant matrix. It is now immediate from Lemma 2.21 that DFT_N is not rigid over a cyclotomic field of order $\tilde{O}(N^3)$. The proof in Theorem 6.2 for DFT_G and G-circulant matrices generalizes directly. $\square$

7 Finite field case

In this section, we sketch how to modify the proofs in the previous sections to deal with matrices over a finite field. The main difficulty that arises when attempting to extend the above methods to finite fields is that the entries of the corresponding DFT matrix might not exist in the field. Furthermore, for a finite field $\mathbb{F}_q$ and integer k with $\gcd(k,q) > 1$, there are no primitive k^{th} roots of unity in any extension of $\mathbb{F}_q$. Because this section involves a significant amount of abstract algebra, we begin by giving a brief overview of the algebraic tools that we will use.

7.1 Preliminaries about Galois theory and finite fields

Our standard reference for field extensions, Galois theory, and finite fields is Chapters V and VI of Lang's Algebra [13]. The monograph by Lidl and Niederreiter [14] is entirely dedicated to finite fields.

Definition 7.1. A field extension $\mathbb{K}/\mathbb{F}$ means that $\mathbb{K}$ is a field and $\mathbb{F}$ is a subfield. The *degree* of the extension is the dimension of $\mathbb{K}$ as a vector space over $\mathbb{F}$. Given a field extension $\mathbb{K}/\mathbb{F}$, we say that $\alpha \in \mathbb{K}$ is *algebraic* over $\mathbb{F}$ if α is a root of some nonzero polynomial P with coefficients in $\mathbb{F}$. We say that $\mathbb{K}/\mathbb{F}$ is an *algebraic extension* if every element of $\mathbb{K}$ is algebraic over $\mathbb{F}$. A *finite extension* is an extension of finite degree.

Fact 7.2. Every finite extension is algebraic.

Definition 7.3. Given a field extension $\mathbb{K}/\mathbb{F}$ and $\alpha \in \mathbb{K}$ that is algebraic over $\mathbb{F}$, we say that the polynomial P over $\mathbb{F}$ is the *minimal polynomial* of α if P is monic and has minimal degree among all nonzero polynomials over $\mathbb{F}$ that have α as a root. The *degree* of α over $\mathbb{F}$ is the degree of its minimal polynomial.

It is not difficult to see that P exists, is unique, and must be irreducible over $\mathbb{F}$.

Fact 7.4. Given a field extension $\mathbb{K}/\mathbb{F}$, let $\alpha \in \mathbb{K}$ be algebraic over $\mathbb{F}$. The set $\mathbb{F}[\alpha]$, defined as the set of all polynomials of α with coefficients in $\mathbb{F}$, is a field. The degree of the extension $\mathbb{F}[\alpha]/\mathbb{F}$ is the degree of α over $\mathbb{F}$.

$\mathbb{F}[\alpha_1, \alpha_2]$ denotes $\mathbb{F}[\alpha_1][\alpha_2]$.

Definition 7.5. Let $\mathbb{K}/\mathbb{F}$ be a field extension and let $\alpha \in \mathbb{K}$ be algebraic over $\mathbb{F}$. Let P be the minimal polynomial of α. The *conjugates* of α are the roots of P (including α itself) in an extension of $\mathbb{K}$ over which P decomposes into linear factors.

Definition 7.6 (Galois extensions)**.** An algebraic field extension $\mathbb{K}/\mathbb{F}$ is *normal* if for every irreducible polynomial P over $\mathbb{F}$, if P has a root in $\mathbb{K}$ then P splits into linear factors over $\mathbb{K}$. The extension $\mathbb{K}/\mathbb{F}$ is *Galois* if it is normal and for all $\alpha \in \mathbb{K}$, all roots of the minimal polynomial of α over $\mathbb{F}$ are distinct.

Fact 7.7 ([13, Ch. V, Thm. 5.5])**.** If $\mathbb{K}$ is a finite field then every extension $\mathbb{K}/\mathbb{F}$ is Galois.

Definition 7.8 (Galois group)**.** For a Galois extension $\mathbb{K}/\mathbb{F}$ we write $\mathrm{Gal}(\mathbb{K}/\mathbb{F})$ to denote the set of those automorphisms of $\mathbb{K}$ that fix $\mathbb{F}$ elementwise.

We begin by stating some basic facts.

Fact 7.9. Let $\mathbb{K}/\mathbb{F}$ be a finite Galois extension of degree g and let $\alpha \in \mathbb{K}$ have degree m. Let $G = \mathrm{Gal}(\mathbb{K}/\mathbb{F})$. Then the following hold.

(i) $|G| = g$.

(ii) $m \mid g$.

(iii) The conjugates of $\alpha \in \mathbb{K}$ are the elements $\pi(\alpha)$ for all $\pi \in \mathrm{Gal}(\mathbb{K}/\mathbb{F})$. The list $(\pi(\alpha) \mid \pi \in G)$ includes each conjugate of α exactly g/m times. In particular, if $\mathbb{K} = \mathbb{F}[\alpha]$ (i. e., $m = g$) then the degree of this extension is the number of conjugates of α.

(iv) $\mathbb{F}$ is precisely the set of common fixed points of $\mathrm{Gal}(\mathbb{K}/\mathbb{F})$.

The following consequence of item (iv) is immediate.

Fact 7.10. Let $\mathbb{K}/\mathbb{F}$ be a Galois extension. Let $\alpha \in \mathbb{K}$ and let $\alpha_1, \dots, \alpha_m$ be the conjugates of α with $\alpha_1 = \alpha$. Then $Q(\alpha_1, \dots, \alpha_m) \in \mathbb{F}$ for any symmetric polynomial Q.

Fact 7.11 ([13, Ch. V, Thm. 5.4])**.** Let $\mathbb{K}$ be a finite field and $\mathbb{F}$ a subfield. If $\mathbb{F} = \mathbb{F}_q$ then $\mathrm{Gal}(\mathbb{K}/\mathbb{F})$ is a cyclic group, generated by the Frobenius automorphism $x \mapsto x^q$.

We have the following consequence.

Fact 7.12. Let $\mathbb{K}$ be a finite field and $\mathbb{F} = \mathbb{F}_q$ a subfield. Let $\alpha \in \mathbb{K}$. Then the conjugates of α over $\mathbb{F}$ are precisely the elements of the form α^{q^j} for nonnegative integers j. In particular, if $\mathbb{K} = \mathbb{F}_{q^m} = \mathbb{F}[\alpha]$ then m is the degree of α over $\mathbb{F}$ and the conjugates of α are exactly $\{\alpha^{q^0}, \alpha^{q^1}, \dots, \alpha^{q^{m-1}}\}$. Moreover, if n is the order of α in the multiplicative group of $\mathbb{K}$ then $m = \mathrm{ord}_q(n)$, the order of q modulo n.

Now we introduce the concept of primitive roots of unity over finite fields and prove some of their basic properties.

Fact 7.13 ([13, Ch. V, Thm. 5.3])**.** The multiplicative group of a finite field is cyclic. (In fact, the finite subgroups of the multiplicative group of any field are cyclic.)

Definition 7.14. Let $\mathbb{F}$ be a field. We say that $\alpha \in \mathbb{F}$ is an n^{th} root of unity if $\alpha^n = 1$. We say that $\alpha \in \mathbb{F}$ is a *primitive n^{th} root of unity* if $\alpha \neq 0$ and the order of α in $\mathbb{F}^\times$ is n.

Fact 7.15 ([14, Thm. 2.47(ii)])**.** A primitive n^{th} root of unity exists in the finite field of order q if and only if $n \mid q-1$.

Fact 7.16. Let $\mathbb{F}$ be a finite field of order q and let $n \mid q-1$. Then the number of n^{th} roots of unity is precisely n, they are the powers of any primitive n^{th} root of unity, and their sum is 0 if $n \geq 2$ and 1 if $n = 1$.

Proof. The n^{th} roots of unity are precisely the roots of the polynomial $x^n - 1$. They form a multiplicative group which is therefore cyclic. The primitive n^{th} roots of unity in $\mathbb{F}$ are precisely the generators of this group. The sum of the roots of $x^n - 1$ is the negative of the coefficient of x^{n-1} in $x^n - 1$. □

Fact 7.17. Let $\mathbb{F}_q$ be the finite field of order q and let n be an integer with $\gcd(q,n) = 1$. Let ω be a primitive n^{th} root of unity in some extension field of $\mathbb{F}_q$. Then the degree of the minimal polynomial of ω over $\mathbb{F}_q$ is $\mathsf{ord}_q(n)$, the order of q modulo n. The conjugates of ω are

$$\omega, \omega^q, \ldots, \omega^{q^{\mathsf{ord}_q(n)-1}}.$$

Proof. Immediate from Fact 7.12. □

7.2 Modifications to the main proofs

In this section, we sketch how to modify the main proofs to work over finite fields. We work over a finite field $\mathbb{F}_q$ where q is a fixed constant (when we say parameters are chosen to be sufficiently large, they may be chosen in terms of q). We will first define the DFT matrices over finite fields.

Definition 7.18. Let $\mathbb{F}_q$ be a finite field and N be an integer with $\gcd(N,q) = 1$. Pick a canonical primitive N^{th} root of unity ω in some extension of $\mathbb{F}_q$. The (x,y) entry of the $N \times N$ matrix $\mathrm{DFT}_{N,\mathbb{F}_q}$ $(0 \leq x, y \leq N-1)$ is ω^{xy}. We will omit the second subscript $\mathbb{F}_q$ and just write DFT_N when the base field is clear from context.

Remark 7.19. The matrix $\mathrm{DFT}_{N,\mathbb{F}_q}$ is well-defined over any extension of $\mathbb{F}_q$ that contains ω. It can easily be verified that the properties proved in Section 2 (namely that DFT_N diagonalizes circulant matrices) also hold in the finite field setting.

The first lemma in this section allows us to lift to a field extension and then argue that if a matrix is highly non-rigid over some low-degree extension then it also cannot be rigid over the base field.

Lemma 7.20. *Consider a finite field $\mathbb{F}_q$ and a finite extension $\mathbb{F}_q[\gamma]$ where $\gamma \neq 0$. If the degree of γ over $\mathbb{F}_q$ is g then for any matrix $M \in \mathbb{F}_q^{n \times n}$ and any positive integer r,*

$$r_M^{\mathbb{F}_q}(gr) \leq r_M^{\mathbb{F}_q[\gamma]}(r).$$

Proof. Let the conjugates of γ be $\gamma_1,\dots,\gamma_g$ where $\gamma_1=\gamma$. Let k be a positive integer such that $\gamma_1^k+\cdots+\gamma_g^k\neq 0$. Such $0\le k\le g-1$ exists because the columns of the Vandermonde matrix generated by the γ_i are linearly independent. Let $s=r_M^{\mathbb{F}_q[\gamma]}(r)$. There must be a matrix $E\in\mathbb{F}_q[\gamma]^{n\times n}$ with at most s nonzero entries in each row and column such that $\mathsf{rank}_{\mathbb{F}_q[\gamma]}(M-E)\le r$. Now consider the g matrices $E_1=E,E_2,\dots,E_g$ where E_i is obtained by taking E and replacing γ with its i^{th} conjugate, γ_i. While naturally, we would like to consider the matrix $(E_1+\cdots+E_g)/g$ and write

$$M-\frac{E_1+\cdots+E_g}{g}=\frac{M-E_1}{g}+\cdots+\frac{M-E_g}{g},$$

the above expression is only valid when $\gcd(g,q)=1$ so we will need a slight modification. Define the matrix E' as follows.

$$E'=\frac{1}{\gamma_1^k+\cdots+\gamma_g^k}\cdot\left(\gamma_1^kE_1+\cdots+\gamma_g^kE_g\right).$$

Note that $\gamma_1^k+\cdots+\gamma_g^k\in\mathbb{F}_q$ and also $\gamma_1^kE_1+\cdots+\gamma_g^kE_g\in\mathbb{F}_q^{n\times n}$ since the entries are symmetric polynomials in $(\gamma_1,\dots,\gamma_g)$. Thus $E'\in\mathbb{F}_q^{n\times n}$ and E' clearly has at most s nonzero entries in each row and column. Next we observe that

$$M-E'=\frac{1}{\gamma_1^k+\cdots+\gamma_g^k}\cdot\left(\gamma_1^k(M-E_1)+\cdots+\gamma_g^k(M-E_g)\right).$$

Note that $\mathsf{rank}_{\mathbb{F}_q[\gamma_i]}(M-E_i)\le r$ for all i. This is because the determinant of every $r\times r$ submatrix of $M-E$ can be written as a formal polynomial in γ with coefficients in $\mathbb{F}_q$ and since γ is a root of each of these polynomials, γ_i must be as well, implying that the determinant of each of the $r\times r$ submatrices of $M-E_i$ is 0. Thus, we conclude that $\mathsf{rank}_{\mathbb{F}_q}(M-E')\le gr$. Writing $M=(M-E')+E'$, we immediately get the desired conclusion. ∎

Following the proof of Theorem 3.1, we can prove an analogue over finite fields. All we needed in Theorem 3.1 was that we were working over a field that contained the roots of unity in the definition of the GWH matrix. Over finite fields, it suffices to work over an extension that contains the necessary roots of unity.

Theorem 7.21. *Let $\mathbb{F}_q$ be a finite field and $N=d^n$ for positive integers d,n,q with $\gcd(d,q)=1$. Let ω be a primitive d^{th} root of unity in some extension of $\mathbb{F}_q$. Let $0<\varepsilon<0.01$ and assume $n\ge 1/\psi$ where*

$$\psi=\frac{\varepsilon^2}{400\log^2(1/\varepsilon)d\log d}.$$

Let $H_{d,n}=\underbrace{\mathrm{DFT}_d\otimes\cdots\otimes\mathrm{DFT}_d}_{n}$. Then

$$r_{H_{d,n}}^{\mathbb{F}_q[\omega]}\left(N^{1-\psi}\right)\le N^{\varepsilon}.$$

With the above, we can now prove a finite field version of Lemma 4.10. The proof is the same as the proof of Lemma 4.10, using Theorem 7.21 in place of Theorem 3.1. The only necessary change is that we need to work over an extension of $\mathbb{F}_q$ that contains all of the necessary roots of unity.

Lemma 7.22. *Let $0 < \varepsilon < 0.01$ be some chosen parameter, $\mathbb{F}_q$ be a fixed finite field, and D be some sufficiently large constant (possibly depending on ε and q). Consider positive integers $t_1 \le t_2 \cdots \le t_n$ with $\gcd(t_i, q) = 1$ for all i. Also assume $a_i \ge \max\left(\frac{t_i^2(\log t_i)^2}{\varepsilon^{10}}, D\right)$ for all i. Let $P = t_1^{a_1} \cdots t_n^{a_n}$ and $L = \lceil 2\log\log P \rceil$. Consider the field extension $\mathbb{F}_q[\omega_1, \dots, \omega_n]$ where ω_i is a primitive $t_i{}^{th}$ root of unity.*[6] *Let DFT_{t_i} be the $t_i \times t_i$ DFT matrix with entries in the field extension. Let*

$$A = \underbrace{\left(\mathrm{DFT}_{t_1} \otimes \cdots \otimes \mathrm{DFT}_{t_1}\right)}_{a_1} \otimes \cdots \otimes \underbrace{\left(\mathrm{DFT}_{t_n} \otimes \cdots \otimes \mathrm{DFT}_{t_n}\right)}_{a_n}.$$

Then we have

$$r_A^{\mathbb{F}_q[\omega_1, \dots, \omega_n]}\left(P^{1-\varepsilon^6/(10Lt_n^2 \log t_n)}\right) \le P^{5\varepsilon}.$$

We also need a slight modification in the proof of Lemma 4.15. We will use the following definitions from Section 4.

- x is a sufficiently large integer.
- l is an integer such that
$$g_{100}(x) \le l \le g_{10}(x)$$
where $g_k(x)$ is defined as in Notation 5.1.
- N is (l, x)-factorable.
- $\mathsf{mult}_N(S), \mathsf{fact}_N(S)$ are defined as in Definition 4.11 and the matrix $M(S)$ is defined as in Definition 4.13.

Remark 7.23. Note that since x is sufficiently large and N is (l, x)-factorable, $\gcd(N, q) = 1$. This will be important later on.

Remark 7.24. Note that if γ is a primitive N^{th} root of unity, the matrix $M(S)$ is defined over the extension $\mathbb{F}_q[\gamma]$ for all subsets S.

Lemma 7.25. *Let $\mathbb{F}_q$ be a fixed finite field. Let x be sufficiently large and $N = q_1 q_2 \cdots q_l$ be an (l, x)-factorable number with $\gcd(N, q) = 1$ and $g_{100}(x) \le l \le g_{10}(x)$. Let $t_1, \dots, t_a$ be the set of prime powers at most $x^{0.3}$ that are relatively prime to q. Let $\omega_1, \dots, \omega_a$ be primitive $t_1{}^{th}, \dots, t_a{}^{th}$ roots of unity and let γ be a primitive N^{th} root of unity. For a subset $S \subset [l]$ with $|S| = k$ and $M(S)$ (as defined in Definition 4.13) a $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ matrix, we have*

$$r_{M(S)}^{\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_a]}\left(\frac{\mathsf{fact}_N(S)}{\exp\left(\varepsilon^6 x^{0.37}\right)}\right) \le \left(\mathsf{fact}_N(S)\right)^{6\varepsilon}$$

as long as $k \ge \frac{x}{(\log x)^{C_0+200}}$.

[6] The ω_i need not be distinct. We only need $\mathbb{F}_q[\omega_1, \dots, \omega_n]$ to be an extension that contains all of $\omega_1, \dots, \omega_n$.

Proof Sketch. Without loss of generality $S = \{1,2,\ldots,k\}$. Recall that in the proof of Lemma 4.15, we argued that the matrix $M(S)$ is $\mathbb{Z}_{q_1-1} \times \cdots \times \mathbb{Z}_{q_k-1}$ circulant. Then, using the prime factorizations of each of $q_1-1,\ldots,q_k-1$, we wrote $\mathbb{Z}_{q_1-1} \times \cdots \times \mathbb{Z}_{q_k-1}$ as a direct product of cyclic groups of prime power order. Since each q_i-1 must factor into prime powers that are at most $x^{0.3}$, we argued that some of these cyclic groups must appear many times in the direct product and then we could apply Lemma 4.10. Over finite fields, the only necessary change in the proof of Lemma 4.15 is due to the fact that for an integer b with $\gcd(b,q)>1$, primitive b^{th} roots of unity do not exist over an extension of $\mathbb{F}_q$. Thus, in the direct product of cyclic groups of prime power order, we cannot use those factors whose order is not relatively prime to q (because we cannot diagonalize $\mathbb{Z}_b$-circulant matrices when $\gcd(b,q)>1$). To deal with this, we will use a more precise bound than (4.1) where prime powers not relatively prime to q are also excluded from the product on the left hand side.

Consider the factorizations of $q_1-1,\ldots,q_k-1$ into prime powers. For each prime power $p_i^{e_i}$ with $p_i^{e_i} \le x^{0.3}$, let $c(p_i^{e_i})$ be the number of indices j for which $p_i^{e_i}$ appears (exactly) in the factorization of q_j-1. Also let p be the characteristic of the finite field $\mathbb{F}_q$ that we are working over (so q is a power of p). Note that

$$(q_1-1)\cdots(q_k-1) = \prod_t t^{c(t)} = p^{c(p)} p^{2c(p^2)} \cdots p^{fc(p^f)} \prod_{gcd(t,p)=1} t^{c(t)},$$

where t ranges over all prime powers at most $x^{0.3}$ and p^f is the largest power of p that is at most $x^{0.3}$. For a power of p, say p^i, let $d(p^i)$ be the number of indices j such that q_j-1 is divisible (not necessarily exactly divisible) by p^i. Let $L = \lfloor (1000+C_0)\log_p \log x \rfloor$. Then we have

$$\begin{aligned} p^{c(p)} p^{2c(p^2)} \cdots p^{fc(p^f)} &= p^{d(p)+d(p^2)+\cdots+d(p^f)} \le p^{\sum_{i=1}^{L} d(p^i) + \sum_{i=L+1}^{f} d(p^i)} \le p^{Lk + fx/(\log x)^{1000+C_0}} \\ &\le (\log x)^{(1000+C_0)k} x^{x/(\log x)^{1000+C_0}}. \end{aligned}$$

Next, consider all prime powers $p_i^{e_i}$ for which $c(p_i^{e_i}) < x^{0.62}$. These satisfy

$$\prod_{t, c(t) \le x^{0.62}} t^{c(t)} \le \left((x^{0.3})^{x^{0.62}} \right)^{x^{0.3}} \le x^{x^{0.92}}.$$

Now without loss of generality, say $\{t_1,\ldots,t_n\}$ is the subset of $\{t_1,\ldots,t_a\}$ ($n \le a$) consisting of the set of prime powers for which $\gcd(t_i,p)=1$ and $c(t_i) \ge x^{0.62}$. Let $P = t_1^{c(t_1)} \cdots t_n^{c(t_n)}$. From the above we know that as long as x is sufficiently large

$$P \ge \frac{\mathsf{fact}_N(S)}{x^{x^{0.92}} (\log x)^{(1000+C_0)k} x^{g_{1000}(x)}} \ge (\mathsf{fact}_N(S))^{(1-\varepsilon)} \cdot \frac{\left(\frac{x}{(\log x)^{C_0+1}}\right)^{\varepsilon k}}{x^{x^{0.92}} (\log x)^{(1000+C_0)k} x^{g_{1000}(x)}} \ge (\mathsf{fact}_N(S))^{(1-\varepsilon)}.$$

Recall $g_{1000}(x) = x/(\log x)^{1000+C_0}$ is as defined in Notation 5.1. The remainder of the proof can be completed in the same way as Lemma 4.15 using Lemma 7.22 in place of Lemma 4.10. □

Using the above we can prove the following analogue of Theorem 4.7.

Theorem 7.26. *Let $\mathbb{F}_q$ be a fixed finite field and $0 < \varepsilon < 0.01$ be some constant. Let x be sufficiently large and $N = q_1 q_2 \cdots q_l$ be an (l,x)-factorable number with $\gcd(N,q) = 1$ and $g_{100}(x) \le l \le g_{10}(x)$. Let $t_1, \dots, t_a$ be the set of prime powers at most $x^{0.3}$ that are relatively prime to q. Let $\omega_1, \dots, \omega_a$ be primitive ${t_1}^{th}, \dots, {t_a}^{th}$ roots of unity and let γ be a primitive N^{th} root of unity. Then*

$$r_{\mathrm{DFT}_N}^{\mathbb{F}_q[\gamma,\omega_1,\dots,\omega_n]}\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.36}\right)}\right) \le N^{7\varepsilon}.$$

Proof. As in the proof of Theorem 4.7, we can subdivide the matrix DFT_N into submatrices of the form $M(S)$ for various subsets $S \subset [l]$ using Lemma 4.14 (it is easily verified that Lemma 4.14 also holds over finite fields). We can remove all of the rows and columns corresponding to integers divisible by too many of the primes $q_1, \dots, q_l$ because the contribution of these rows and columns is low-rank. The remaining entries can be subdivided into matrices of the form $M(S)$ where $|S|$ is sufficiently large so we can then apply Lemma 7.25 to change a small number of entries in each row and column to reduce the rank significantly. The precise computations are exactly the same as in Theorem 4.7. □

We will now combine Theorem 7.26 with Lemma 7.20 to get our main theorem for circulant matrices over finite fields.

Theorem 7.27. *Let $0 < \varepsilon < 0.01$ be a given parameter and $\mathbb{F}_q$ be a fixed finite field. For all sufficiently large N, if M is an $N \times N$ circulant or Toeplitz matrix then*

$$r_M^{\mathbb{F}_q}\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.35}\right)}\right) \le N^{15\varepsilon}.$$

Proof. First we analyze circulant matrices of size N_0 where N_0 is (l,x)-factorable for some

$$g_{100}(x) \le l \le g_{10}(x).$$

Note that as long as x is sufficiently large, N_0 must be relatively prime to q. Since DFT matrices diagonalize circulant matrices (even over finite fields), Theorem 7.26 and Lemma 2.21 imply that for M_0, an $N_0 \times N_0$ circulant matrix where N_0 satisfies the previously mentioned conditions,

$$r_{M_0}^{\mathbb{F}_q[\gamma,\omega_1,\dots,\omega_a]}\left(\frac{2N_0}{\exp\left(\varepsilon^6(\log N_0)^{0.36}\right)}\right) \le N_0^{14\varepsilon}$$

where γ is a primitive N^{th} root of unity and $\omega_1, \dots, \omega_a$ are primitive ${t_1}^{\mathrm{th}}, \dots, {t_a}^{\mathrm{th}}$ roots of unity for $t_1, \dots, t_a$ being the set of prime powers at most $x^{0.3}$ that are relatively prime to q. Now we analyze the degree of the extension $\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_a]$. Note $\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_a] \subset \mathbb{F}_q[\eta]$ where η is a primitive root of unity of order $C = N_0 \operatorname{lcm}(t_1, t_2, \dots t_a)$. By Fact 7.17, the degree of the extension $\mathbb{F}_q[\eta]$ is the order of q modulo C. Since N_0 factors into a product of distinct x-good primes, by Fermat's little theorem, the order of q modulo N_0 divides $\left(x^{0.3}\right)!$. Also, all prime powers dividing $\operatorname{lcm}(t_1, t_2, \dots, t_a)$ are at most $x^{0.3}$. Thus, the order of q modulo C divides $\left(x^{0.3}\right)!$. Overall, the order of $q \mod C$ is at most

$$\left(x^{0.3}\right)! < x^{0.3x^{0.3}} \le \exp\left((\log N_0)^{0.31}\right).$$

Thus the degree of the extension $\mathbb{F}_q[\gamma,\omega_1,\dots,\omega_a]$ is at most $\exp\left((\log N_0)^{0.31}\right)$. By Lemma 7.20

$$\mathsf{r}^{\mathbb{F}_q}_{M_0}\left(\frac{N_0}{\exp\left(\varepsilon^6(\log N_0)^{0.359}\right)}\right) \le N_0^{14\varepsilon}.$$

To complete the proof, we can simply repeat the arguments in the proof of Theorem 5.3. For a circulant matrix M of arbitrary size $N\times N$, note that it is possible to embed an M in the upper left corner of a circulant matrix of any size at least $2N$. By Lemma 5.2, there exists an N_0 that is (l,x)-factorable for some $g_{100}(x) \le l \le g_{10}(x)$ such that

$$\frac{N_0}{(\log N_0)^2} \le N \le \frac{N_0}{2}.$$

We deduce

$$\mathsf{r}^{\mathbb{F}_q}_{M}\left(\frac{N_0}{\exp\left(\varepsilon^6(\log N_0)^{0.359}\right)}\right) \le N_0^{14\varepsilon}.$$

Rewriting the bounds in terms of N we get

$$\mathsf{r}^{\mathbb{F}_q}_{M}\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.35}\right)}\right) \le N^{15\varepsilon}. \qquad \square$$

8 G-circulant matrices over finite fields

We will now generalize Theorem 6.2 to matrices over a finite field $\mathbb{F}_q$ except we will require the additional condition that $\gcd(|G|,q)=1$. Write the underlying abelian group G as a direct product of cyclic groups $\mathbb{Z}_{n_1}\times\cdots\times\mathbb{Z}_{n_a}$. While for matrices with entries in $\mathbb{C}$, it sufficed to work with the Kronecker product of the DFT matrices $\mathrm{DFT}_{n_1}\otimes\cdots\otimes\mathrm{DFT}_{n_a}$, we require slightly different techniques for rigidity over a fixed finite field as an extension containing all of the necessary roots of unity could have too high degree. Instead of working through DFT matrices, we will work directly with the G-circulant matrices themselves.

While for sufficiently large cyclic groups, we did not require the condition that $\gcd(|G|,q)=1$ (see Theorem 7.27), we require the condition for general abelian groups because we need to use Theorem 7.21 to deal with the case when G contains the direct product of many copies of a small cyclic group. In particular, our techniques do not handle a group such as $\mathbb{Z}_{p^2}\times\cdots\times\mathbb{Z}_{p^2}$ where p is equal to the characteristic of the field $\mathbb{F}_q$. It is an interesting open question to see if the condition that $\gcd(|G|,q)=1$ can be eliminated. The work in [7] deals with the case where $q=p^a$ for a prime p and G is a direct product of many cyclic groups of order p but not the case when G is a direct product of many cyclic groups of order p^2 (or some other power of p).

The first important observation is that Theorem 5.3 can be slightly strengthened so that to reduce the rank of any circulant matrices, the locations to be changed are fixed and the changes are fixed linear combinations of the entries of the circulant matrix. More precisely, we make the following definition.

Definition 8.1. Given a group G of order $|G|=n$, we say G is (r,s)-reducible over $\mathbb{F}_q$ if the following condition holds. There exist

- a set $S\subset[n]\times[n]$ of positions where S contains at most s positions in each row and column,

- matrices $A, B \in \mathbb{F}_q^{n\times n}$ where $\mathsf{rank}(A), \mathsf{rank}(B) \le r$,
- matrices $E_1, \dots, E_n \in \mathbb{F}_q^{n\times n}$ with all nonzero entries in S, and
- matrices $Y_1, \dots, Y_n, Z_1, \dots, Z_n \in \mathbb{F}_q^{n\times n}$

such that for any G-circulant matrix M with top row $(x_1, \dots, x_n)$, we have

$$M = A(x_1Y_1 + \cdots + x_nY_n) + (x_1Z_1 + \cdots + x_nZ_n)B + (x_1E_1 + \cdots + x_nE_n).$$

In such a decomposition, the matrices A, B will be called (r,s)-reduction matrices and the matrices $Y_1, \dots, Y_n, Z_1, \dots, Z_n, E_1, \dots, E_n$ will be called (r,s)-reduction helpers. We write $Y_M = x_1Y_1 + \cdots + x_nY_n$ and similarly for Z_M and E_M.

Following the proof of Theorem 5.3, we can show that $\mathbb{Z}_N$ is

$$\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.35}\right)}, N^{15\varepsilon}\right)$$

reducible over $\mathbb{C}$. We now prove an analogue of this result for finite fields.

Claim 8.2. *For fixed* $0 < \varepsilon < 0.01$ *and all sufficiently large N, the group* $\mathbb{Z}_N$ *is*

$$\left(\frac{N}{\exp\left(\varepsilon^6(\log N)^{0.35}\right)}, N^{15\varepsilon}\right)$$

reducible over $\mathbb{F}_q$.

Proof. First consider an integer N_0 that is (l,x)-factorable for some $g_{100}(x) \le l \le g_{10}(x)$. As long as x is sufficiently large, $\gcd(N_0, q) = 1$. Let M_0 be a $N_0 \times N_0$ circulant matrix (i. e., a G-circulant matrix for $G = \mathbb{Z}_{N_0}$) over $\mathbb{F}_q$ and let the entries in its top row be $x_1, \dots, x_{N_0}$. Let γ be a primitive N_0^{th} root of unity and $t_1, \dots t_n$ be the set of prime powers at most $x^{0.3}$ that are relatively prime to q. Let $\omega_1, \dots, \omega_n$ be roots of unity of order $t_1, \dots, t_n$ respectively. By Theorem 7.26, there exists a matrix E over $\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_n]$ with at most $N_0^{7\varepsilon}$ nonzero entries in each row and column such that

$$\mathsf{rank}(\mathrm{DFT}_{N_0} - E) \le \frac{N_0}{\exp\left(\varepsilon^6(\log N_0)^{0.36}\right)}.$$

Now write

$$M_0 = \mathrm{DFT}_{N_0}^* \cdot D \cdot \mathrm{DFT}_{N_0} = (\mathrm{DFT}_{N_0} - E)^* D \cdot \mathrm{DFT}_{N_0} + E^* D(\mathrm{DFT}_{N_0} - E) + E^* DE$$

where D is a diagonal matrix whose entries are linear combinations of $x_1, \dots, x_{N_0}$. Note that all of the above matrices have entries contained in $\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_n] \subseteq \mathbb{F}_q[\eta]$ where η is a primitive root of unity of order $C = N_0 \operatorname{lcm}(t_1, \dots, t_n)$. As argued before in the proof of Theorem 7.27, the degree of the extension is at most $\exp\left((\log N_0)^{0.31}\right)$. Let the conjugates of η be $\eta_1 = \eta, \eta_2, \dots, \eta_m$. Let $\mathrm{DFT}_{N_0}^1, \dots, \mathrm{DFT}_{N_0}^m$ be obtained by taking DFT_{N_0} and replacing η with its conjugates. Define $D^1,\dots,D^m,E^1,\dots,E^m$ similarly. As in the proof of Lemma 7.20, there exists an integer k such that $\eta_1^k+\cdots+\eta_m^k\neq 0$. We now have

$$M_0=\frac{1}{\eta_1^k+\cdots+\eta_m^k}\left(\sum_{i=1}^m \eta_i^k(\mathrm{DFT}_{N_0}^i-E^i)^*D^i\cdot\mathrm{DFT}_{N_0}^i+\sum_{i=1}^m \eta_i^k {E^i}^* D^i(\mathrm{DFT}_{N_0}^i-E^i)+\sum_{i=1}^m \eta_i^k {E^i}^* D^i E^i\right).$$

Note that $1/(\eta_1^k+\cdots+\eta_m^k)\in\mathbb{F}_q$ and all three of the sums are matrices whose entries are linear combinations of $x_1,\dots,x_{N_0}$ with coefficients in $\mathbb{F}_q$. The last term satisfies the desired sparsity constraint as it has at most $N_0^{14\varepsilon}$ nonzero entries in each row and column and the locations of these entries are independent of M_0.

It remains to argue that the first two terms satisfy the desired rank constraint. Note that the span of the columns of $(\mathrm{DFT}_{N_0}^1-E^1),\dots,(\mathrm{DFT}_{N_0}^m-E^m)$ has dimension at most

$$\frac{mN_0}{\exp\left(\varepsilon^6(\log N_0)^{0.36}\right)}\leq\frac{N_0}{\exp\left(\varepsilon^6(\log N_0)^{0.359}\right)}$$

over $\mathbb{F}_q[\eta]^{N_0}$. Therefore, the dimension of the intersection of this subspace with $\mathbb{F}_q^{N_0}$, say V, has dimension at most

$$\frac{N_0}{\exp\left(\varepsilon^6(\log N_0)^{0.359}\right)}.$$

In particular we can write

$$\sum_{i=1}^m \eta_i^k(\mathrm{DFT}_{N_0}^i-E^i)^*D^i\cdot\mathrm{DFT}_{N_0}^i=x_1C_1+\cdots+x_{N_0}C_{N_0}$$

for some fixed matrices $C_1,\dots,C_{N_0}$ with entries in $\mathbb{F}_q$. Also all columns of $C_1,\dots,C_{N_0}$ must be in V so each can be written as AY_i where A is a fixed matrix with rank at most

$$\frac{mN_0}{\exp\left(\varepsilon^6(\log N_0)^{0.36}\right)}.$$

Thus there exists fixed matrices $Y_1,\dots,Y_{N_0}\in\mathbb{F}_q^{n\times n}$ and a matrix A satisfying the desired rank constraint such that

$$\sum_{i=1}^m \eta_i^k(\mathrm{DFT}_{N_0}^i-E^i)^*D^i\cdot\mathrm{DFT}_{N_0}^i=A(x_1Y_1+\cdots+x_{N_0}Y_{N_0}).$$

A similar argument shows that the second term can also be written in the desired form.

Now to extend to arbitrary N (not necessarily (l,x)-factorable), simply note that any circulant matrix of size N can be embedded into a circulant matrix of any given size at least $2N$ where each entry of the larger matrix is equal to some entry of the original matrix. We can then apply Lemma 5.2 and complete the proof in the same way as Theorem 5.3. □

Claim 8.2 allows us to deal with large cyclic groups. We will also need a way of dealing with a direct product of many copies of a small cyclic group.

Claim 8.3. *Let $\mathbb{F}_q$ be a finite field and $N = d^n$ for positive integers d,n,q with $\gcd(d,q) = 1$. Let $0 < \varepsilon < 0.01$ and assume $n \geq 2/\psi$ where*

$$\psi = \frac{\varepsilon^2}{400\log^2(1/\varepsilon)d\log d}.$$

Let $G = \underbrace{\mathbb{Z}_d \otimes \cdots \otimes \mathbb{Z}_d}_{n}$. Then G is $\left(N^{1-\psi/2}, N^{2\varepsilon}\right)$ reducible over $\mathbb{F}_q$.

Proof. Let ω be a primitive d^{th} root of unity in some extension of $\mathbb{F}_q$. We use Theorem 7.21 to find a sparse matrix E with entries in $\mathbb{F}_q[\omega]$ such that $H_{d,n} - E$ has low rank where

$$H_{d,n} = \underbrace{\mathrm{DFT}_d \otimes \cdots \otimes \mathrm{DFT}_d}_{n}.$$

We can then repeat the same argument as in the proof of Claim 8.2, using the fact that $H_{d,n}$ diagonalizes any G-circulant matrix. $\square$

Now we introduce the main technical result of this section that allows us to deal with direct products of different groups without going through the corresponding DFT matrices.

Claim 8.4. *Consider a list of abelian groups, $G_1, \dots, G_a$, such that $|G_i| = n_i$. Assume for each $1 \leq i \leq a$, G_i is (r_i, s_i)-reducible over $\mathbb{F}_q$. Let $G = G_1 \times \cdots \times G_a$ and $|G| = n = n_1 n_2 \cdots n_a$. Then the group G is (r,s)-reducible over $\mathbb{F}_q$ where*

$$r = \sum_{S \subset [a], |S| = l} 2^l \prod_{i \in S} \sqrt{r_i n_i} \prod_{i' \notin S} n_{i'},$$

$$s = \sum_{S \subset [a], |S| < l} 2^{|S|} \prod_{i \in S} n_i \prod_{i' \notin S} s_{i'}.$$

Proof. Let M be a G-circulant matrix. For each $1 \leq i \leq a$, let A^i, B^i be the (r_i, s_i)-reduction matrices for the group G_i. Note that the reducibility assumption means that any G_i-circulant matrix can be written as a sum of three matrices where the first contains a fixed high dimensional subspace in its left nullspace, the second contains a fixed high dimensional subspace in its right nullspace, and the third is sparse. The first step in our proof will involve writing M as a sum of 3^a matrices. Roughly, each of these 3^a matrices corresponds to choosing one of the three possible components (large left nullspace, large right nullspace, or sparse) for each of the groups G_i.

More formally, for each $i \in [a]$, consider the group G_i. Let its (r_i, s_i)-reduction helpers be $\{Y_{g_i}\}$, $\{Z_{g_i}\}$, $\{E_{g_i}\}$ for $g_i \in G_i$. Let $Y(g_i) = A^i Y_{g_i}$, $Z(g_i) = Z_{g_i} B^i$ and $E(g_i) = E_{g_i}$. By definition, for a G-circulant matrix M_{G_i} with top row given by $\{x_g\}$ for $g \in G_i$,

$$M_{G_i} = \sum_{g_i \in G} (Y(g_i) + Z(g_i) + E(g_i)) x_{g_i}.$$

Thus, for fixed $g_i, h_i, k_i \in G_i$, the entry of $Y(g_i) + Z(g_i) + E(g_i)$ indexed by (h_i, k_i) is equal to 1 if $h_i + k_i = g_i$ (in the group G_i) and 0 otherwise.

We index the rows and columns of M with ordered tuples $(h_1,\dots,h_a)$ and $(k_1,\dots,k_a)$ respectively (where $h_i,k_i \in G_i$). Let the entries in the top row of M be $x_{g_1,\dots,g_a}$ where $(g_1,\dots,g_a)$ ranges over $G_1\times\cdots\times G_a$. Now for each ordered tuple $I=(i_1,\dots,i_a)\in\{1,2,3\}^a$, we will construct a $|G|\times|G|$ matrix M_I. Let $S^1(I),S^2(I),S^3(I)\subset[a]$ denote the subsets of locations where the entry of I is $1,2$ or 3 respectively. We define

$$M_I=\sum_{(g_1,\dots,g_a)} x_{g_1,\dots,g_a}\left(\bigotimes_{i\in S^1(I)} Y(g_i)\right)\otimes\left(\bigotimes_{i\in S^2(I)} Z(g_i)\right)\otimes\left(\bigotimes_{i\in S^3(I)} E(g_i)\right)$$

where the sum is over all $(g_1,\dots,g_a)\in G_1\times\cdots\times G_a$. The first important observation is that

$$M=\sum_{I\in\{1,2,3\}^a} M_I\,. \tag{8.1}$$

To see this, it suffices to note that the coefficients of $x_{g_1,\dots,g_a}$ on the right hand side for fixed $g_1,\dots,g_a$ are given by the matrix

$$\sum_{I\in\{1,2,3\}^a}\left(\bigotimes_{i\in S^1(I)} Y(g_i)\right)\otimes\left(\bigotimes_{i\in S^2(I)} Z(g_i)\right)\otimes\left(\bigotimes_{i\in S^3(I)} E(g_i)\right)= \bigotimes_{i\in[a]}\left(Y(g_i)+Z(g_i)+E(g_i)\right).$$

The entry indexed by $(h_1,\dots,h_a)$ and $(k_1,\dots,k_a)$ on the right hand side is equal to 1 if $h_i+k_i=g_i$ for all i and 0 otherwise. This completes the proof of (8.1).

We would like to write M as a sum of three matrices, say P_1,P_2,P_3, whose entries are linear forms in the variables $x_{g_1,\dots,g_a}$ and such that $P_1=AY, P_2=ZB$ for some fixed low-rank matrices A,B and P_3 is sparse. Write

$$M=\sum_{\substack{I\in\{1,2,3\}^a\\ |S^3(I)|\le a-l}} M_I+\sum_{\substack{I\in\{1,2,3\}^a\\ |S^3(I)|> a-l}} M_I\,.$$

We will prove that P_1,P_2 can be obtained by splitting the first sum and we can set P_3 to be equal to the second sum. For each $1\le i\le a$, there exists a set of linearly independent vectors $v_1^i,\dots,v_{n_i-r_i}^i$ such that $v_j^iA^i=0$ and a set of linearly independent vectors $u_1^i,\dots,u_{n_i-r_i}^i$ such that $B^iu_j^i=0$ for all $1\le j\le n_i-r_i$. We can complete the set $\{v_1^i,\dots,v_{n_i-r_i}^i\}$ to a basis $\{v_1^i,\dots,v_{n_i}^i\}$ and similar for $\{u_1^i,\dots,u_{n_i}^i\}$. Consider the basis of $\mathbb{F}_q^n$ consisting of the vectors $v_{j_1}^1\otimes v_{j_2}^2\otimes\cdots\otimes v_{j_a}^a$ where $(j_1,\dots,j_a)\in[n_1]\times\cdots\times[n_a]$. Now assume we are given a matrix M_I with $I\in\{1,2,3\}^a$. The key observation is that if $j_i\le n_i-r_i$ for some index $i\in S^1(I)$, then

$$\left(v_{j_1}^1\otimes v_{j_2}^2\otimes\cdots\otimes v_{j_a}^a\right)M_I=0\,. \tag{8.2}$$

This is because $Y(g_i)=A^iY_{g_i}$ so by construction, $v_{j_i}^iY(g_i)=0$ for all $g_i\in G_i$. Using this observation and examining the definition of M_I, we immediately get (8.2). Similarly, we get that if $j_i\le n_i-r_i$ for some index $i\in S^2(I)$ then

$$M_I\left(u_{j_1}^1\otimes u_{j_2}^2\otimes\cdots\otimes u_{j_a}^a\right)=0\,.$$

Let $R_1 \subset \{1,2,3\}^a$ be the set of ordered tuples I such that

$$\prod_{i \in S^1(I)} \frac{r_i}{n_i} \leq \prod_{i \in S^2(I)} \frac{r_i}{n_i}\,.$$

Let $R_2 = \{1,2,3\}^a \backslash R_1$. We now write

$$\sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| \leq a-l}} M_I = \sum_{\substack{I \in R_1 \\ |S^3(I)| \leq a-l}} M_I + \sum_{\substack{I \in R_2 \\ |S^3(I)| \leq a-l}} M_I$$

and will argue that the first term, which we call P_1, has a fixed, high dimensional subspace contained in its left nullspace while the second term, which we call P_2, has a fixed, high dimensional subspace contained in its right nullspace. We work with the basis $v^1_{j_1} \otimes v^2_{j_2} \otimes \cdots \otimes v^a_{j_a}$ where $(j_1, \dots, j_a) \in [n_1] \times \cdots \times [n_a]$ and count the number of these basis vectors that are not in the left nullspace of P_1. For each $I \in R_1$, by (8.2), M_I contributes at most

$$\prod_{i \in S^1(I)} r_i \prod_{i \in [a] \backslash S^1(I)} n_i$$

basis vectors for which vM_I is nonzero. Furthermore if $S^1(I) \subset S^1(I')$ for two distinct ordered tuples I and I', the contributions of $M_{I'}$ are redundant with the contributions of M_I. Thus, we can ignore the contributions of M_I for ordered tuples I for which $|S^3(I)| < a-l$. Overall, the number of basis vectors outside the left nullspace of P_1 is at most

$$\sum_{\substack{I \in R_1 \\ |S^3(I)| = a-l}} \prod_{i \in S^1(I)} r_i \prod_{i \in [a] \backslash S^1(I)} n_i \leq \sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| = a-l}} \prod_{i \in [a] \backslash S^3(I)} \sqrt{r_i n_i} \prod_{i \in S^3(I)} n_i = \sum_{S \subset [a], |S| = l} 2^l \prod_{i \in S} \sqrt{r_i n_i} \prod_{i' \notin S} n_{i'} \tag{8.3}$$

where to obtain the above inequality, we first used the fact that $I \in R_1$ and then used that for a fixed set $S^3(I)$, there are 2^l possible ordered tuples I. Note that the basis $v^1_{j_1} \otimes v^2_{j_2} \otimes \cdots \otimes v^a_{j_a}$ where $(j_1, \dots, j_a) \in [n_1] \times \cdots \times [n_a]$ is fixed (i. e., independent of the entries of M). Thus we can write $P_1 = AX$ where the entries of X are linear forms in the entries of M and A is a fixed matrix with rank bounded above by the expression in (8.3). A similar argument allows us to write $P_2 = YB$ for a fixed matrix B with the desired rank.

Now it remains to bound the sparsity of

$$\sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| > a-l}} M_I\,.$$

We claim that the number of nonzero entries in each row and column of M_I is at most

$$\prod_{i \in S^3(I)} s_i \prod_{i \in [a] \backslash S^3(I)} n_i\,.$$

To see this, note that for each i, the matrices $E(g_i)$, as g_i ranges over all of G_i, have all of their nonzeros contained in a fixed set S_i where S_i contains at most s_i distinct locations in each row and column. For each fixed subset $S^3(I)$, there are exactly $2^{|S^3(I)|}$ possible ordered tuples I. Thus the number of nonzero entries in each row and column of the sum is at most

$$\sum_{\substack{I\in\{1,2,3\}^a \\ |S^3(I)|>a-l}} \prod_{i\in S^3(I)} s_i \prod_{i\in[a]\setminus S^3(I)} n_i = \sum_{S\subset[a],|S|<l} 2^{|S|} \prod_{i\in[a]\setminus S} s_i \prod_{i\in S} n_i \,.$$

This completes the proof that the group G is (r,s)-reducible over $\mathbb{F}_q$. □

We are now ready to prove the main theorem about rigidity of G-circulant matrices over finite fields.

Theorem 8.5. *Let $\mathbb{F}_q$ be a fixed finite field and $\varepsilon < 0.01$ be a fixed constant. Let G be an abelian group. As long as $|G|$ is sufficiently large and $\gcd(|G|,q)=1$, for any G-circulant matrix M over $\mathbb{F}_q$, we have*

$$r_M^{\mathbb{F}_q}\left(\frac{|G|}{\exp\left(\varepsilon^{20}(\log|G|)^{0.3}\right)}\right) \le |G|^{100\varepsilon}.$$

Proof. By the Fundamental Theorem of Finite Abelian Groups we can write $G=\mathbb{Z}_{n_1}\times\cdots\times\mathbb{Z}_{n_a}$. The proof will essentially follow the same method as the proof of Theorem 6.1 except using Claim 8.4 to deal with direct products of cyclic groups that are roughly the same order.

Without loss of generality, $n_1\le n_2\le\cdots\le n_a$. We will choose k to be a fixed, sufficiently large positive integer (possibly depending on q,ε). Consider the ranges $I_1=[k,k^2), I_2=[k^2,k^4),\ldots I_j=[k^{2^{j-1}},k^{2^j})\ldots$ and so on. Let S_j be a multiset defined by $S_j = I_j\cap\{n_1,\ldots,n_a\}$. Fix a j and let the elements of S_j be $x_1\le\cdots\le x_b$. By Claim 8.2, we have that (since k sufficiently large) for each x_i, the group $\mathbb{Z}_{x_i}$ is

$$\left(\frac{x_i}{\exp\left(\varepsilon^6(\log x_i)^{0.35}\right)}, x_i^{15\varepsilon}\right) \tag{8.4}$$

reducible over $\mathbb{F}_q$. Now we will use Claim 8.4 to argue about the group $G_j=\mathbb{Z}_{x_1}\times\cdots\times\mathbb{Z}_{x_b}$. Set $l=\lceil\varepsilon b\rceil$ in Claim 8.4. We get that $G_j=\mathbb{Z}_{x_1}\times\cdots\times\mathbb{Z}_{x_b}$ is (r,s) reducible for some r,s that are obtained by plugging (8.4) into the expressions in Claim 8.4. We bound r and s more carefully below. We have

$$\begin{aligned} r &\le \binom{b}{\lceil\varepsilon b\rceil}\frac{2^{\lceil\varepsilon b\rceil}x_1\cdots x_b}{\left(\exp\left(\varepsilon^6(\log x_1)^{0.35}\right)\right)^{\lceil\varepsilon b\rceil/2}} \le \frac{(2b)^{\lceil\varepsilon b\rceil}}{(\frac{\varepsilon b}{3})^{\lceil\varepsilon b\rceil}}\frac{x_1\cdots x_b}{\left(\exp\left(\varepsilon^6(\log x_1)^{0.35}\right)\right)^{\lceil\varepsilon b\rceil/2}} \\ &= x_1\cdots x_b\left(\frac{36}{\varepsilon^2\exp\left(\varepsilon^6(\log x_1)^{0.35}\right)}\right)^{\lceil\varepsilon b\rceil/2}. \end{aligned}$$

As long as k is sufficiently large, we have

$$\begin{aligned} r &\le x_1\cdots x_b\left(\frac{36}{\varepsilon^2\exp\left(\varepsilon^6(\log x_1)^{0.35}\right)}\right)^{\lceil\varepsilon b\rceil/2} \le x_1\cdots x_b\left(\frac{1}{\exp\left(\varepsilon^6(\log x_1)^{0.34}\right)}\right)^{\lceil\varepsilon b\rceil/2} \\ &\le \frac{x_1\cdots x_b}{\exp\left(\varepsilon^7(\log x_1\cdots x_b)^{0.33}\right)} \end{aligned}$$

where in the last step we used the fact that $x_i\le x_1^2$ for all i. Next we bound the sparsity s. We have

$$s\le 4^b x_b\cdots x_{b-\lfloor\varepsilon b\rfloor+1}(x_{b-\lfloor\varepsilon b\rfloor}\cdots x_1)^{15\varepsilon} = 4^b(x_1\cdots x_b)^{15\varepsilon}(x_b\cdots x_{b-\lfloor\varepsilon b\rfloor+1})^{1-15\varepsilon} \le (x_1\cdots x_b)^{18\varepsilon}$$

where in the last step above, we used the fact that $x_i \le x_1^2$ for all i. Thus, we have shown that the group $G_j = \mathbb{Z}_{x_1} \times \cdots \times \mathbb{Z}_{x_b}$ is

$$\left(\frac{x_1 \cdots x_b}{\exp\left(\varepsilon^7 (\log x_1 \cdots x_b)^{0.33}\right)}, (x_1 \cdots x_b)^{18\varepsilon}\right)$$

reducible over $\mathbb{F}_q$.

The above allows us to deal with direct products of large cyclic groups that are all of roughly the same order. In the direct product $G = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_a}$, we will split the terms into cyclic groups of small order, which can be dealt with using Claim 8.3, and several products of large cyclic groups that can each be dealt with using the above. We now formalize this argument. For each integer c between 2 and k with $\gcd(c,q) = 1$, let m_c be the number of copies of c in the set $\{n_1, \ldots, n_a\}$. If $m_c \ge k^2(\log k)^2/\varepsilon^4$ then by Claim 8.3, the group $\underbrace{\mathbb{Z}_c \times \cdots \times \mathbb{Z}_c}_{m_c}$ is

$$\left(c^{m_c(1-\varepsilon^4/(k^2(\log k)^2)}, c^{2m_c\varepsilon}\right)$$

reducible over $\mathbb{F}_q$. Let $L = \lceil 2 \log\log |G| \rceil$ and ensure that $|G|$ is sufficiently large so that $L > k$. Let T be the set of integers c between 2 and k with $\gcd(c,q) = 1$ such that $c^{m_c} \ge |G|^{\varepsilon/(2L)}$. Note that as long as $|G|$ is sufficiently large, all elements of T must satisfy $m_c \ge k^2(\log k)^2/\varepsilon^4$. Let R be the set of indices j for which $\prod_{x \in S_j} x \ge |G|^{\varepsilon/(2L)}$. Note that S_j is clearly empty for $j \ge L$. Recall that $\gcd(|G|,q) = 1$ so the group G can be written as

$$G = \left(\bigtimes_{\substack{2 \le c < k \\ \gcd(c,q)=1}} \left(\underbrace{\mathbb{Z}_c \times \cdots \times \mathbb{Z}_c}_{m_c}\right)\right) \times \left(\bigtimes_{1 \le j \le L} G_j\right).$$

Define

$$B = \left(\bigtimes_{c \notin T} \left(\underbrace{\mathbb{Z}_c \times \cdots \times \mathbb{Z}_c}_{m_c}\right)\right) \times \left(\bigotimes_{j \notin R} G_j\right).$$

Note that

$$|B| \le \left(|G|^{\varepsilon/(2L)}\right)^{k+L} \le |G|^{\varepsilon}.$$

Also $G = B \times D$ where

$$D = \left(\bigtimes_{c \in T} \left(\underbrace{\mathbb{Z}_c \otimes \cdots \otimes \mathbb{Z}_c}_{m_c}\right)\right) \times \left(\bigtimes_{j \in R} G_j\right).$$

Now we apply Claim 8.4 again on D where we view D as a direct product of groups of the form

$\underbrace{\mathbb{Z}_c \otimes \cdots \otimes \mathbb{Z}_c}_{m_c}$ and G_j and we set $l = 1$. We get that D is (r_D, s_D) reducible over $\mathbb{F}_q$ where

$$r_D \leq 2|D| \left(\sum_{c \in T} \frac{1}{c^{m_c \varepsilon^4/(2k^2 \log k)}} + \sum_{j \in R} \frac{1}{\exp\left(0.5\varepsilon^7 (\log \prod_{x \in S_j} x)^{0.33}\right)} \right) \leq \frac{|D|}{\exp(\varepsilon^8 (\log |G|)^{0.32})},$$
$$s_D \leq |D|^{18\varepsilon}.$$

Finally, note that the group B is trivially $(0, |B|)$ reducible over $\mathbb{F}_q$. Thus, by Claim 8.4, $G = B \times D$ is (r_G, s_G) reducible over $\mathbb{F}_q$ for

$$r_G \leq \frac{2|D| \cdot |B|}{\exp(0.5\varepsilon^8 (\log |G|)^{0.32})} \leq \frac{|G|}{\exp(\varepsilon^8 (\log |G|)^{0.31})},$$
$$s_G \leq |B| \cdot |D|^{18\varepsilon} \leq |G|^{19\varepsilon}.$$

The above immediately implies that for any G-circulant matrix M with $\gcd(|G|, q) = 1$,

$$\mathsf{r}_M^{\mathbb{F}_q}\left(\frac{|G|}{\exp(\varepsilon^{20} (\log |G|)^{0.3})}\right) \leq |G|^{100\varepsilon}.$$

This completes the proof.

□

9 Final remarks and open questions

Our main results, Theorems 6.2, 6.3, and 8.5, naturally raise some open questions. Recall that the $N \times N$ DFT (Discrete Fourier Transform) matrix is the matrix (ω^{ij}) $(i, j = 0, \ldots, N-1)$ where ω is a primitive N^{th} root of unity.

- Are the $N \times N$ DFT matrices rigid over the N^{th} cyclotomic field $\mathbb{Q}[\omega]$ (where ω is a primitive N^{th} root of unity) ? (Compare this question with Theorem 6.3.)
- Do there exist circulant matrices, or G-circulant matrices for some class of abelian groups G, that are rigid over $\mathbb{Q}$? (Again, compare with Theorem 6.3.)
- Does there exist a finite field $\mathbb{F}_q$ and G-circulant matrices for some class of abelian groups G with $\gcd(|G|, q) > 1$ that are rigid over $\mathbb{F}_q$? (Compare this question with Theorem 8.5.)
- Do there exist rigid G-circulant matrices over $\mathbb{C}$ for some class of (necessarily non-abelian) groups G ?

When G is non-abelian, it is no longer possible to simultaneously diagonalize the matrices $M_G(f)$ for all f but there is a change of basis matrix A such that $AM_G(f)A^*$ is block-diagonal where the diagonal blocks correspond to the irreducible representations of G. When all of the irreducible representations of G have small degree (dimension), it may be possible to use similar techniques to the ones used here.

On the other hand, this suggests that perhaps $M_G(f)$ is a candidate for rigidity when all irreducible representations of G have large degree. Frobenius proved in 1896 that the group $SL_2(\mathbb{F}_p)$ of 2×2 matrices over $\mathbb{F}_p$ with determinant 1 has no nontrivial irreducible representations of degree less than $(p-1)/2$ over $\mathbb{C}$ and thus has highly nonabelian structure [10]. (See [6] for an accessible presentation.) Thus we make the following conjecture.

Conjecture 9.1. *For large primes p, a random G-circulant* $(0,1)$*-matrix* $M_G(f)$ *for* $G = SL_2(\mathbb{F}_p)$ *is Valiant-rigid over* $\mathbb{C}$ *with high probability. Here by "random" we mean the function* $f : SL_2(\mathbb{F}_p) \to \{0,1\}$ *is chosen randomly.*

Acknowledgments

We thank Lajos Rónyai for suggesting the literature references cited in Section 7.1. We would like to thank Laci Babai and the editorial team at ToC for many important comments.

References

[1] JOSH ALMAN AND LIJIE CHEN: Efficient construction of rigid matrices using an NP oracle. In *Proc. 60th FOCS*, pp. 1034–1055. IEEE Comp. Soc., 2019. [doi:10.1109/FOCS.2019.00067] 3

[2] JOSH ALMAN AND RYAN WILLIAMS: Probabilistic rank and matrix rigidity. In *Proc. 49th STOC*, pp. 17:1–17:23. ACM Press, 2017. [doi:10.1145/3055399.3055484, arXiv:1611.05558] 3, 4, 6, 8

[3] RICHARD ARRATIA AND LOUIS GORDON: Tutorial on large deviations for the binomial distribution. *Bull. Mathematical Biology*, 51(1):125–131, 1989. [doi:10.1007/BF02458840] 16

[4] LÁSZLÓ BABAI AND BOHDAN KIVVA: Matrix rigidity: More conjectures refuted. In preparation. 5

[5] ROGER C. BAKER AND GLYN HARMAN: Shifted primes without large prime factors. *Acta Arithmetica*, 83(4):331–361, 1998. [doi:10.4064/aa-83-4-331-361] 7, 17

[6] GIULIANA DAVIDOFF, PETER SARNAK, AND ALAIN VALETTE: *Elementary Number Theory, Group Theory, and Ramanujan Graphs*. Volume 55 of *LMS Student Texts*. Cambridge Univ. Press, 2003. [doi:10.1017/CBO9780511615825] 46

[7] ZEEV DVIR AND BENJAMIN EDELMAN: Matrix rigidity and the Croot-Lev-Pach lemma. *Theory of Computing*, 15(8):1–7, 2019. [doi:10.4086/toc.2019.v015a008, arXiv:1708.01646] 3, 4, 6, 8, 14, 37

[8] ZEEV DVIR AND ALLEN LIU: Fourier and circulant matrices are not rigid. In *34th Computational Complexity Conference (CCC 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019. [doi:10.4230/LIPIcs.CCC.2019.17]

[9] JOEL FRIEDMAN: A note on matrix rigidity. *Combinatorica*, 13(2):235–239, 1993. [doi:10.1007/BF01303207] 3

[10] FERDINAND GEORG FROBENIUS: Über Gruppencharaktere. *Sitzungsberichte der Preußischen Akademie der Wissenschaften zu Berlin*, 39:985–1021, 1896. 46

[11] ODED GOLDREICH AND AVISHAY TAL: Matrix rigidity of random Toeplitz matrices. *Comput. Complexity*, 27(2):305–350, 2018. Preliminary version in STOC'16. [doi:10.1007/s00037-016-0144-9] 3, 7

[12] ABHINAV KUMAR, SATYANARAYANA V. LOKAM, VIJAY M. PATANKAR, AND M. N. JAYALAL SARMA: Using elimination theory to construct rigid matrices. *Comput. Complexity*, 23(4):531–563, 2014. [doi:10.1007/s00037-013-0061-0, arXiv:0910.5301] 3

[13] SERGE LANG: *Algebra*. Volume 211 of *Grad. Texts in Math.* Springer, 3rd edition, 1996. 30, 31

[14] RUDOLF LIDL AND HARALD NIEDERREITER: *Finite Fields*. Encycl. Math. Appl. Cambridge Univ. Press, 2nd edition, 1996. [doi:10.1017/CBO9780511525926] 30, 32

[15] SATYANARAYANA V. LOKAM: On the rigidity of Vandermonde matrices. *Theoret. Comput. Sci.*, 237(1–2):477–483, 2000. [doi:10.1016/S0304-3975(00)00008-6] 3

[16] SATYANARAYANA V. LOKAM: Quadratic lower bounds on matrix rigidity. In *Internat. Conf. on Theory and Appl. of Models of Computation (TAMC'06)*, pp. 295–307. Springer, 2006. [doi:10.1007/11750321_28] 3

[17] SATYANARAYANA V. LOKAM: Complexity lower bounds using linear algebra. *Foundations and Trends in Theoretical Computer Science*, 4(1–2):1–155, 2009. [doi:10.1561/0400000011] 3

[18] MOHAMMAD AMIN SHOKROLLAHI, DANIEL A. SPIELMAN, AND VOLKER STEMANN: A remark on matrix rigidity. *Inform. Process. Lett.*, 64(6):283–285, 1997. [doi:10.1016/S0020-0190(97)00190-7] 3

[19] LESLIE G. VALIANT: Graph-theoretic arguments in low-level complexity. In *Math. Found. Comp. Sci. (MFCS'77)*, pp. 162–176. Springer, 1977. [doi:10.1007/3-540-08353-7_135] 3, 11

AUTHORS

Zeev Dvir
Associate professor
Department of Mathematics and
Department of Computer Science
Princeton University
Princeton, NJ, USA
zdvir@princeton.edu
https://www.cs.princeton.edu/~zdvir

Allen Liu
Student
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, MA, USA
cliu568@mit.edu

ABOUT THE AUTHORS

ZEEV DVIR was born in Jerusalem, Israel. He received his Ph. D. from the Weizmann Institute in Israel in 2008. His advisors were Ran Raz and Amir Shpilka. He has a broad interest in theoretical computer science and mathematics and especially in computational complexity, pseudorandomness, coding theory and discrete mathematics.

ALLEN LIU was born in Rochester, NY. He is currently a fourth-year undergraduate student in mathematics at the Massachusetts Institute of Technology. His research interests include computational complexity and learning theory.